

Issues Surrounding the Long-Run Performance of Initial Public Offerings

Eric Brown

Ph. D. thesis

The University of Edinburgh

2003



Declaration

This thesis has been composed entirely by the author. All work is attributable to and any errors are the responsibility of the author. None of the work contained in the thesis has been submitted for any other degree or qualification.

Signed

Eric Brown

Table of Contents

Chapter 1: Introduction	1
Chapter 2 : Long-Run Event Study Methodology	
2.1 Introduction	6
2.2 Measurement of Abnormal Return	
2.2.1 What is the Normal return for a Stock?	7
2.2.2 Aggregation and Choice of Statistic	9
2.2.3 Equal Weighted or Value Weighted Averages	14
2.3 Potential Biases and Sources of Misspecification	17
2.3.1 Known Problems with Hypothesis Tests on Long-Run Stock Returns	18
2.3.2 Calendar Time Techniques Revisited	26
2.4 Conclusion	32
Chapter 3 : Long-Run Performance Analysis of a New Sample of IPOs	
3.1 Introduction	33
3.2 Data	36
3.3 Methodology	39
3.4 Results	45
3.4.1 Initial Returns and the "Hot Markets Phenomenon"	45
3.4.2 Long-Run Abnormal Returns	49
3.5 Conclusion	57
Chapter 4 : Further Features of Long-Run EPO Performance	
4.1 Introduction	59
4.2 Existing Theoretical Research	59
4.3 Decomposition of Returns	66
4.3.1 Data	67
4.3.2 Methodology	69
4.3.3 Results	72
Chapter 5 : A Learning-Based Model with Unknown Expected Return	
5.1 Introduction	77
5.2 The Model	78
5.2.2 Predictability of returns and Excess Volatility of Prices and Returns	89
5.3 Unknown Expected Return	91
5.4 Conclusions	101
Chapter 6: Conclusion	103
Appendix A1 : Gauss programme for Skew Adjusted t-Test Procedures	105
Appendix A2 : Gauss Programme for Simulation results of Chapter 5	108
Appendix A3 : Derivation of Variance of p_t and $\hat{\beta}_t^*$	111
References	113

Chapter 1. Introduction

Loosely, the Efficient Markets Hypothesis (EMH) is the notion that all available information is properly reflected in prevailing prices¹. The key implication, (just as loosely), is that future prices are unpredictable on the basis of currently available information. In fact, since Samuelson set out his famous “Proof that Properly Anticipated Prices Fluctuate Randomly” (Samuelson, (1965)), it has become clear that prices can exhibit a *certain type* of predictability even in a fully efficient market. This may occur, for example, when expected returns are time varying.

The volume of research on the efficiency of financial markets is quite staggering. This is natural since the hypothesis is a fundamental tenet of much mainstream financial economics and the large financial markets with their standardized contracts, official exchanges etc, and with the enormous economic resources which are directed towards them, appear one of the most likely areas of the economy in which to identify informational efficiency. However the task has never been straightforward. One reason is that a formal representation of EMH inevitably involves two separate hypotheses: one of these is rational expectations, the other is a model of price (or returns) determination. This complicates the interpretation of empirical work on EMH since if a predictable pattern in the data is identified it would appear at first glance that a supernormal profit has been available to anyone taking advantage of this predictability. However we can only ever say that the profit was supernormal with reference to a particular model of returns determination and the possibility always exists that the wrong model has been employed.

¹ Fama sets out precisely what can be implied by the EMH in the famous survey Fama (1970)

Another reason for the vast body of research is that a position on EMH is not so much a binary choice as a matter of degree: it is likely that large established financial markets are generally highly efficient but surely significant inefficiencies have arisen in the past and others will arise in the future, particularly in smaller or developing markets. As such the empirical question is less one of “Are markets efficient?” than “Where do notable market inefficiencies exist?” The scope for research naturally becomes much broader.

We might also add that the EMH is inherently an intriguing hypothesis: most economists would be delighted if their research led them to discover predictable profit opportunities in patterns of observed asset prices!

The research in this thesis is *not* intended to represent a critique on the EMH but the hypothesis is always there in the background, part of the general context of the issues discussed². We are interested in whether observed asset pricing is rational and we focus on Initial Public Offerings (IPOs), stocks newly offered for sale on official stock exchanges. It has been observed by Ritter (1991) *inter alia* that IPOs appear to display behaviour which is difficult to reconcile with the EMH, specifically that on average they jump up in price during their first day of trading on a stock market (*positive* abnormal performance), and furthermore, from this point onwards they exhibit *negative* abnormal performance over a period of years. This is a typical example of what has been described in the literature as an “EMH anomaly”. Other such anomalies include the January effect, where it has been noted that stocks generally perform abnormally well in January, and the Closed-End Fund Puzzle whereby investment trusts³ (or

² The critique on EMH, together with a comprehensive survey of the so-called “EMH anomalies” literature, is available in Fama (1998).

³ An investment trust is an investment vehicle which takes the form of a listed company whose only assets are shares in other companies.

closed-end funds, as they are known in USA) tend to trade in the market at a price which is less than their Net Asset Value.

These and other patterns have been highlighted over the last 30 years or more and the observations have provoked several branches of research in response. In one direction some writers have disputed results, taking issue with econometric techniques employed. There are several key components in empirical work on EMH, one of which obviously relates to the dual hypothesis problem mentioned above: how should abnormal returns be measured properly? There can be further difficulties associated with aggregation and hypothesis testing and we go into this subject in some detail in Chapter 2. We focus on a particular type of investigation, the long-run event study, in which the research involves taking a sample of firms which have experienced an event such as an IPO and tracking their performance from the time of the event in order to see whether the event has had an effect on their long-run performance. We discuss the various methodological advances which have been made in this area of econometrics and we try to provide an up-to-date guide on best practice for this type of research.

Another line of research into the EMH anomalies has been to test the robustness of existing results by undertaking parallel studies using data from different markets or different time periods. Again, interpretation of results can be difficult: if it appears that an anomaly no longer exists the explanation may be that once the predictable profit opportunity came to light practitioners began trading in the market (profitably) so as to eliminate it or, alternatively, the earlier results may have arisen through faulty econometric methodology or pure chance. This drawback notwithstanding, in Chapter 3 we carry out a new empirical study on the long-run performance of IPOs issued to UK stock markets during the mid-1990s. We employ as many as possible of the best

techniques identified in Chapter 2. The results of our investigation are mixed: although the initial positive abnormal performance identified by Ritter (1991) is clearly present in data set at hand, we find little evidence of long-run underperformance except during an unconventional extension in which we attempt to control for industry sector in the measurement of abnormal returns.

Although we cannot claim clear-cut confirmation of long-run IPO underperformance from our empirical results we turn in the second half of the thesis to theoretical work which might inform on the subject. In Chapter 4 we survey the theories which have been advanced and we discuss learning-based models in particular. Learning based explanations of long-run underperformance are attractive because the fundamental premise that less is known about an IPO than is known about an ordinary seems very reasonable. More importantly, learning-based explanations are perhaps unique in being able to capture the process through which a stock evolves from being an (underperforming) IPO into, at some point, an (ordinarily priced) stock. One such model, due to Morris (1996), motivates further empirical work, the results of which are also presented in Chapter 4. In this section we try to decompose long-run abnormal performance into an abnormal dividend performance and an abnormal price performance.

Finally, in Chapter 5, we develop an alternative learning based model of asset pricing in which agents learn about the value of a less stylized asset than that which is analysed in Morris (1996). We note that this model shares several features with another model by Lewellen and Shanken (2000) which applies the learning concept to a different area of financial economics. We show how our own model can be viewed as an extension of the simplest version of LS (2000) in which agents have an extra parameter (expected return) about which to learn. In the model expected return is driven by the Capital Asset

Pricing Model (CAPM) and we use numerical simulations to see whether agents might overestimate the CAPM beta and hence overprice the security during the learning process. Such a bias is indeed evident but the simulation results are disappointing in that we also find that Beta cannot be estimated with a satisfactory degree of accuracy using the procedure implied by the model.

Taken together the various pieces of work presented here are not conclusive on the question of long-run underperformance of IPOs. Like other efficient markets anomalies long-run IPO underperformance may have existed in some markets at certain points in time, and may be in the process of disappearing. It is hoped, however, that as well as providing new evidence on the specific question of long-run IPO performance, the research in this thesis illustrates the various routes which can be taken to investigate this type of issue, and highlights the difficulties which can arise.

Chapter 2. Long-Run Event Study Methodology

2.1. Introduction

Research on long-run IPO underperformance, (as well as the effects on long-run performance attributed to stock-splits, rights issues, seasoned equity offerings etc) casts doubt on the Efficient Markets Hypothesis (EMH), but as always with EMH two explanations are possible: the empirical observations may indeed represent violations of EMH, or, alternatively, they may be the result of faulty econometric methodology/data mining/pure chance. In order to illustrate the likelihood of the latter explanation, this chapter explains the methodological difficulties which arise when measuring and testing the significance of long-run abnormal security returns. Section 2.2 is concerned with how properly to measure abnormal returns. Section 2.3 considers the testing of the hypothesis that abnormal returns are zero. Many sources of test misspecification are identified. Section 2.4 briefly concludes.

2.2 Measurement of abnormal returns

In order to measure abnormal return it is necessary to have a benchmark for normal return, or expected return. This is an extremely difficult matter - it is arguably the fundamental question of financial economics. Loosely, the expected return on an asset is determined by its riskiness, but how, precisely, is riskiness determined? Certainly, some proxy for expected return is required in the calculation of abnormal return, (defined as observed return minus expected return.) Having arrived at a value for abnormal return for each observation in a sample, two further questions arise: how should these observations be

aggregated, and how should significance be tested? Each of these questions are discussed in this section.

2.2.1 What is the normal return for a stock?

This of course is a vast subject within finance and financial economics but two models which have proved popular in event studies are the Capital Asset Pricing Model, (CAPM), developed by Sharpe (1964) and generalised by Black (1972), and the Fama and French 3 Factor Model, (FF), due to Fama and French (1992). There are various ways to implement these models in an event study concerned with long-run stock performance. The models themselves state that excess return on an asset or portfolio is determined as follows:

$$\text{CAPM: } r_{it} - r_{rf,t} = \beta_i(r_{mt} - r_{rf,t}) + \varepsilon_{it}$$

$$\text{FF: } r_{it} - r_{rf,t} = \beta_{1i}(r_{mt} - r_{rf,t}) + \beta_{2i}HML_t + \beta_{3i}SMB_t + e_{it}$$

where $r_{it}, r_{mt}, r_{rf,t}$, are, respectively, the return on the asset or portfolio, the return on the market portfolio, and the risk free rate of return at time t , and where HML and SMB are variables reflecting the returns on high book to market value stocks relative to low book to market value stocks, and the returns on small stocks relative to large stocks. The β parameters can be estimated using out of sample data in which case, under the (questionable) assumption of parameter constancy, the errors in the equations above yield estimates of abnormal return. Alternatively the models can be implemented as OLS regressions with intercept in which case the intercept can be interpreted as an estimate of per period abnormal return. Estimation of the β parameters using out of sample data raises a problem in an investigation of IPO returns if it is desirable to use a recent data set. This

technique imposes an extra data requirement (2 years might be a reasonable minimum timespan) for the estimation of model parameters. In an event study investigating the effects of stock splits, for example, this may be less of a problem because data from *before* the event date can be used to estimate the model parameters but in the case of IPOs, such data obviously doesn't exist.

An alternative to these equilibrium models of security returns is the “market adjusted returns model”. In this model the difference between the security return and the return on the market portfolio yields an estimate of abnormal return, ar_{it} :

$$ar_{it} = r_{it} - r_{mt}$$

This can be viewed as the CAPM with the (strictly inaccurate) restriction that $\beta=1$ for every security. Although this model is clearly simplistic, and literally wrong (as are FF and CAPM of course), it nevertheless has some merits. Firstly it has the advantage of not requiring any extra data for parameter estimation. Secondly, Section 2.3 below contains a discussion of the pitfalls and biases which can arise during an event study and in some cases statistics based on the market adjusted returns model are the least misspecified.

The last general approach to the measurement of abnormal returns is conceptually quite different and has gained popularity in recent years: the control firms approach. This approach involves selecting a matching firm/ matching portfolio for each sample firm and taking the difference between sample firm performance and the performance of the matching firm or portfolio as the estimate of abnormal return:

$$ar_{it} = r_{it} - r_{cf,t}$$

The matching firm should control for as many potential determinants of expected returns as possible in order that the abnormal return figures isolate and capture only the effects of the event. Obviously there is something arbitrary about the factors for which the matching firm/portfolio control, and controlling for size will probably yield a different abnormal return from controlling for size and the ratio of book to market value. On the other hand, the estimates may not be *very* different and there is now good evidence, discussed in Section 3, to suggest that test statistics based on this approach remain relatively well specified, in the sense that theoretical size levels prove relatively accurate.

2.2.2 Aggregation and Choice of statistic

Having settled on a model for normal returns, there are several ways of proceeding to estimate abnormal returns over the long-run. The choice is important as a different statistics will not generally preserve the ranking of abnormal returns within a sample. Until very recently this debate concentrated mainly on the choice between the buy and hold abnormal return, (\overline{BHAR}), and the cumulative average abnormal return, (CAAR). As the name suggests, the \overline{BHAR} takes account of the compounding of gains over a long period in a way that the cumulative abnormal return (CAAR) does not. This may or may not be an advantage. From the statistical point of view the distribution of the \overline{BHAR} is usually more non-standard but it has been argued that if this problem is recognized and confronted (often using bootstrapping techniques), the \overline{BHAR} may have superior statistical properties (in the sense that it can be used in relatively high powered accurately sized hypothesis tests). However this is not an uncontested conclusion and in any case very recently the debate has

taken a new direction with some authors arguing that neither the \overline{BHAR} nor the CAAR is appropriate as each fails to take account of cross-correlation of returns within a sample. Instead, Mitchell and Stafford (2000), and Brav, Geczy and Gompers (2000) have argued forcefully that a calendar time application of the Fama and French three factor model, or the calendar time abnormal return (CTAR) should be employed. These concepts are explained carefully in the present section.

2.2.2a The cumulative average abnormal return (CAAR)

The average of the abnormal returns of all firms in a sample in any particular event time period t^i is the average abnormal return (AAR):

$$AAR_t = \frac{1}{n} \sum_{i=1}^n ar_{it}$$

The T-period cumulative average abnormal return, (CAAR), is just the sum of the AARs from the T different event time periods.

$$CAAR_T = \sum_{t=1}^T AAR_t$$

The CAAR can be used to test the hypothesis that the mean monthly abnormal return is zero. If there is no abnormal performance as hypothesized under the null, the CAAR will follow a random walk. An important point is that portfolio rebalancing is implied in the construction of the CAAR. The CAAR is the sum of a succession AARs which are all

¹ An event time period is for example the first month or the second month after issuance rather than say July 1995 which is a calendar time period.

themselves *equal weighted* averages. It doesn't matter how the stock has performed in previous periods, the latest AAR is always calculated on the basis of equal weights. In practice, if one tried to replicate the performance indicated by a CAAR, one would have to sell some of each of the "winners" and buy more of the "losers" at the end of each period in order to get the "portfolio" back to equal weights for the next period. This is essentially a trading strategy and its effect is arguably just to introduce more noise.

2.2.2b The buy and hold abnormal return

As the name suggests this measure includes the effects of compounding. The T-period buy and hold abnormal return for any particular firm in a sample is:

$$BHAR_{iT} = \prod_{t=1}^T (1 + r_{it}) - \prod_{t=1}^T (1 + E(r_t)) .$$

The T-period average buy and hold abnormal return (BHAR) across an entire sample is just the average of the individual T-period BHARs²:

$$\overline{BHAR}_T = \frac{1}{n} \sum_{i=1}^n BHAR_{it}$$

Security returns are positively skewed at all frequencies but 3 year *BHAR*s or 5 year *BHAR*s are extremely positively skewed: This has implications for test specification discussed in Section 3.

² In fact the average can be equal weighted or value weighted (by market capitalization at the time of the event). The choice is important and the issues are discussed in Section 2.2.3 below.

It can be seen that *BHARs* of less than -100% are quite possible, as are 36 month cumulative abnormal returns (CARs) for individual stocks. This possibility, however, need not be a serious concern: a *gross* return of less than -100% on a limited liability security such as a stock would clearly be problematic, but *BHARs* and the other statistics discussed here measure relative performance and a relative return of less than -100% is not necessarily misleading³.

\overline{BHAR} s and CAARs measure different things. Three year \overline{BHAR} s allow tests of the hypothesis that mean abnormal three year returns are zero. Three year CAARs allow tests of the hypothesis that mean monthly abnormal returns during the three year test period are zero. CAARs are biased estimators of \overline{BHAR} s.

2.2.2c Calendar time techniques

\overline{BHAR} s and CAARs are essentially sample means and when we test hypotheses concerning means we rely on the assumption of independence of observations within the sample. If instead the individual abnormal returns in a sample of events are correlated the normal procedure for estimating the standard error of a mean (and the t-statistic) is invalid. It is beyond any argument that corporate events such as equity issuance cluster through time and also by industry and it may be but does not necessarily follow from this that the

³ A statistic which captures the compounding property inherent in the *BHAR* but which can virtually never be less than -1 is the Abnormal Performance Index, (API):

$$ar_{iT} = \prod_{t=1}^T (1 + ar_{it}) - 1, \quad API_T = \frac{1}{n} \sum_{i=1}^n ar_{iT}$$

abnormal returns of equity issuers are correlated. Mitchell and Stafford (2000) claim that the abnormal returns of the constituents of frequently explored samples of event firms are indeed cross-correlated⁴. In Section 2.3.2 we discuss techniques which have been developed to adjust the standard *t*-statistic to account for cross correlation. Alternatively the issue can be confronted directly by forming calendar time portfolios for use in a procedure such as the Calendar Time 3 Factor Regression. This procedure involves, for each calendar month, forming a portfolio which is comprised of all sample firms which have experienced the event within the last (say) 3 years. For each calendar month the portfolio return is calculated and the portfolio excess return (net of the treasury bill rate) becomes the dependent variable in a 3 factor time-series regression:

$$R_{p,t} - R_{f,t} = a + b(R_{m,t} - R_{f,t}) + cSMB_t + dHML_t + e_t$$

where $R_{m,t}$ is the return on the market portfolio as proxied by a broad-based share index and SMB and HML are the Fama and French size and book-to-market value factors described in Section 2.2.1 above. Assuming that the model is the correct model of expected returns the intercept provides a measure of monthly abnormal performance and its standard error is not unduly influenced by any cross-correlation of sample firm returns. The regression is obviously heteroskedastic since the number of firms in the event portfolio changes each month but this can be handled using standard techniques⁵.

⁴ Specifically the data which are investigated by Mitchell and Stafford are samples of firms which have undergone the following corporate events: seasoned equity offerings (SEOs), take-overs and share repurchase scheme announcements.

⁵ Heteroskedasticity is an insignificant problem if as in some implementations only the returns on those calendar months in which the event portfolio contains at least 10 firms are included in procedure.

Another technique which works directly with calendar time portfolios is the Cumulative Calendar Time Abnormal Return (the CCTAR) in which calendar time portfolios are calculated as above but the returns of interest are the abnormal returns as opposed to the raw returns used in the regression above. Abnormal returns for each sample firm can be calculated using any of the procedures described in Section 2.2.1 above and for each calendar month the calendar time portfolio abnormal return is calculated as the average abnormal return of the sample firms included in the portfolio on that particular month. Under the null hypothesis of no abnormal returns the calendar time abnormal return for any month is expected to be zero and the CCTAR which is the cumulative sum of these calendar monthly abnormal returns should behave in the same way as the CAAR described in Section 2.2.1 above, ie it should follow a random walk.

2.2.3 Equal weighted or value weighted averages

\overline{BHAR} s can be constructed as equal weighted averages or they can be value weighted (by initial market capitalisation). Likewise in the calendar time procedures the calendar time monthly portfolios can be constructed in either of these two ways⁶. Equal weighted statistics have been far more popular in the literature but there are some reasons to be more interested in the value weighted statistics. The basic effect of value weighting is clearly to place more emphasis on the performance of the larger firms in a sample. Most well known market indices and sector indices are value weighted as is, of course, the “market portfolio” of CAPM and other equilibrium models.

⁶ Value weighted CAARs have never to the author’s knowledge been reported in published work on financial events such as IPOs. In fact a value weighted CAAR amounts to something extremely similar to but less intuitive than the ordinary buy and hold return (\overline{BHAR}).

Typically, long-run event studies are used to investigate some perceived anomaly in security pricing. If anomalous pricing is confirmed, and the study is credible, results place a question mark over the Efficient Markets Hypothesis itself. Efficient Markets, however, is a hypothesis which is never taken absolutely literally. The common understanding is that real markets are efficient enough such that efficiency is a better approximation than an alternative assumption. Few would argue that violations of market efficiency do not arise. They arise more frequently in smaller markets and in less liquid markets. Within a particular market, they arise more often in connection with smaller less liquid securities. Fama,(1998), has argued that value weighting the observations in an event study will bring to results the appropriate sense of perspective. It is true that weird and inexplicable share price behaviour for a set of very small firms seems less of a reason to discard the market efficiency hypothesis, than if some of the major stocks are displaying sustained systematic abnormal performance. Such small firms will not have as much influence on the realizations of value weighted CAARs and \overline{BHAR} s.

On the other hand a sample of event firms often consists of a large number of very small firms and at the other end perhaps one or two huge firms (this is particularly the case in studies of IPO performance). Value weighting the observations leads to much higher variances for the mean statistics (much less accurate estimates of abnormal performance) as the small number of large firms dominate the sample. The principle cost of value-weighting is therefore that hypothesis tests have lower power.

A study by Brav and Gompers (1997) reinvestigates the long-run performance of IPOs in the US and emphasizes Fama's argument in favour of value weighted statistics. Brav and Gompers are primarily interested in the effects on performance of venture capital backing at the time of the offering. (They suspect that venture capitalist backing may provide easier

access to capital, may impose superior management structures / financial discipline on the firm, and if the market underestimates the value of these functions, venture capital backed IPOs may experience long-run performance which is significantly better than non-venture capital backed IPOs). They compare their full sample results with those in the seminal study of IPO performance by Ritter (1991). They report that even though there are minor differences between the samples, they find underperformance almost identical to Ritter when using equal weighted statistics, but this underperformance largely disappears when they switch to value weights. This is not just down to value weighted statistics having higher standard errors: the point estimate of underperformance is much closer to zero using value weighted \overline{BHAR} s suggesting that in the samples used by Ritter (1991) and Brav and Gompers(1997) it is the small firms which have delivered the negative abnormal performance. (Brav and Gompers also find support for their hypothesis that venture capital backing leads to superior long-run performance.)

One pragmatic conclusion may be that researchers should present both value-weighted and equal-weighted results but as in the choice between CAARs and \overline{BHAR} s, the exact hypothesis under investigation may often suggest which form of statistic should be emphasized. For example if we are interested in estimating the abnormal performance of a randomly selected firm around the time of some particular event then the equal weighted averages are more insightful, whereas if the context is the debate on the empirical validity of the Efficient Markets Hypothesis then value weighting may well be more appropriate.

A final point suggests that the separate questions discussed in this section may not be independent: some authors have pointed out that certain models of expected return appear to be much more successful at pricing some stocks than others. In particular Brav, Geczy and Gompers (2000) have shown that the Fama and French 3 factor model is less

successful at pricing small firms (ie if the returns on a portfolio of small firms are used in the regression outlined in 2.2.1 the intercept is unlikely to be close to zero). As such equal weighted portfolios may be more appropriate in an implementation of the Fama and French model since the “bad model problem”⁷ will be less severe.

2.3. Potential biases and sources of misspecification

As discussed above there is no perfect model of expected returns, so abnormal return will always be measured with error. Nor, usually, is it possible to estimate the variance of normal returns with confidence – estimates tend to be quite model specific. Over the last twenty years, these problems and others have gradually come to light in a series of papers on event study methodology. Important papers have often generated results by simulating event studies. Most of the findings, and some of the simulation methods, are explained in Section 2.3.1. Section 2.3.2 focuses specifically on what is currently the really contentious issue, the issue of cross correlation of sample returns. Section 2.4, the Conclusion, tries to draw together the results from this literature in order to make prescriptions for long-run event studies in the future. The exposition draws on research by Brown and Warner (1980, 1985), Barber and Lyon (1997), Kothari and Warner (1997), Lyon Barber and Tsai (1999), Brav and Gompers (1997), Mitchell and Stafford (2000), Loughran and Ritter (2000), and Brav Geczy and Gompers (2000) (BW, BK, KW, LBT, BG, MS, LR and BGG respectively), eclectically and not always in chronological order.

⁷ This phrase, emphasizing that all models are literally wrong, was coined by Fama (1991)

2.3.1 Known problems with hypothesis tests on long-run stock returns

Brown and Warner (1980) take 250 different samples of 50 securities from the set of all firms with monthly data available from CRSP⁸. They select event dates randomly and then collect data from event month -89 to event month 10. Their aim is to introduce (artificial) abnormal performance of 1% or 5% in event month zero, and then study the power of various tests to detect this abnormal performance, conditional on a particular model of abnormal returns. They hope, thereby, to assess the efficacy of various different models of normal returns.

In fact some peripheral results of the Brown and Warner simulations are particularly important: even before introducing abnormal performance, there appear to be significant differences between the empirical sampling distribution of test statistics and the distribution which is assumed for the hypothesis tests. Under one of the models of expected return they find that the t-tests reject the null of no abnormal performance in favour of negative abnormal returns twice as often as it should. They conclude that: “tests for abnormal security price performance can be misleading and must be interpreted with great caution”⁹.

Brown and Warner (1980, 1985) try to identify good procedures by focussing on the power of the resulting test statistics. In the process they uncovered several potential problems with the true size of such tests. Kothari and Warner, (1997) and Barber and Lyon (1997) have really scrutinised the extent of and explanations for misspecification of tests based on CAARs and *BHAR*s in event studies.

⁸ Centre for Research into Securities Prices, Chicago, the principal source of financial data for academic research in US.

Kothari and Warner construct 250 samples of 200 stocks picked randomly with replacement from the population of all securities having any returns data during a time period. Once a security has been selected in this way, an event month is selected randomly. The data requirement is then that returns data must be available for months -24 to 0 (so that model parameters can be estimated)¹⁰. They note firstly that the CAARs on these samples have positive mean, ie the CAARs are biased.

They suggest several possible reasons for this bias: a rebalancing bias, a data requirements bias, and a skewness bias.

The rebalancing bias is connected with the portfolio rebalancing implicit in the calculation of CAARs and outlined in section 2.2.1 above. Blume and Stambaugh (1983) noted that nonsynchronous trading and/or bid ask bounce effects can cause returns calculated on this type of rebalanced portfolio to be biased. To see this clearly, imagine that the real price of a stock doesn't move for many periods on end, but the published closing price is either the offer price or the bid price depending on whether the customer involved in the last trade was a buyer or a seller, then the CAAR calculation will include a spurious (unattainable) trading profit arising from the assumed rebalancing. Naturally this effect becomes more significant the bigger is the proportional bid-offer spread and the shorter is the time period.

⁹ BW(1980, page 218)

¹⁰These data requirements induce a survivorship bias, explained below and recognized by the writers

Kothari and Warner suggest that if this rebalancing is the source of bias in CAARs, we should expect the mean of sample \overline{BHAR} s to be zero (since there is less rebalancing¹¹). Unfortunately, it is not. It too is positive¹². (table 3, p315)

The contribution of skewness to the misspecification of tests is best illustrated by Barber and Lyon (1997) who describe a very simple experiment: take a large number of samples of 100 observations from a χ -square distribution with one degree of freedom, and on the basis of each sample, carry out a t-test to test the hypothesis that the population mean is one in a two tailed test at the 5% level of significance. (The skewness of the Chi-square (1) is comparable to the skewness of monthly stock returns). The null will be rejected in over 6% of the samples, with the vast majority of rejections resulting from observations falling in the left tail of the distribution. The reason for this bias is that there is positive correlation between the sample mean and the sample standard deviation of a skewed distribution: when the sample mean is high, it is usually because the sample includes outliers from the positive tail of the distribution. These outliers boost the sample standard deviation and lower the t statistic making it harder to reject the null. This is an important general result: the distribution of stock returns is always skewed, more so as the time period gets longer. When the distribution of abnormal returns on stocks is likewise skewed, ordinary t-tests are inherently biased towards detecting negative abnormal performance. The central limit theorem suggests that large enough samples will eliminate this bias. This has been

¹¹ there will still be some rebalancing bias in the market portfolio return as long as a stock index rather than a control firm is used, because the index itself is rebalanced at certain times

¹² Kothari and Warner (1997) find that buy and hold returns have fractionally *higher* (positive) mean than CAARs, yet Barber and Lyon, using an apparently very similar population and sampling procedure, find that the mean \overline{BHAR} is negative as Kothari and Warner predicted. The writers of both papers remark on this difference, and it illustrates how fragile are some of the findings from this type of investigation into test specification

confirmed empirically by Lyon, Barber and Tsai (1998), but in practice such a large sample may not be available¹³.

KW employ an ingenious test to demonstrate that conditioning a sample on prior data availability can raise its return. They take a date, January 1980 as it happens, and they measure the monthly returns on all securities that existed on the CRSP tapes on that date. Then they measure the monthly returns of these same securities (if they still exist) over each of the following 11 months. Then they calculate the grand mean of this huge sample.

Next they repeat the process except that this time they condition their sample on 1 year's prior data availability, (ie they exclude any observation for which the firm did not exist on the tapes one year earlier). They recalculate this grand mean *over the same time period*, and they find that as the sample size falls from 317709 to 296341, the grand mean return rises. The grand mean rises (and the sample size falls) monotonically as the sample is conditioned on 2, 3, and 4 years of prior data availability. Conditioning on 4 years data reduces the number of observations to 247856 and raises the mean monthly return by 0.2%. This appears to suggest that during their sample period, newly listed firms, (essentially IPOs), performed poorly.

The general existence of a new listings bias remains an open question and is discussed fully in the next chapter which contains empirical research on the long-run performance of IPOs. KW's results from this particular experiment however should be treated with a certain amount of caution: the measured returns are raw returns, (unadjusted for risk). In addition, some of the new issues which dragged down returns performance would have been closed-end funds which are normally excluded from samples in studies which measure IPO

¹³ Considering monthly returns, Lyon, Barber, and Tsai found inaccurate size in t-tests on sample

abnormal performance¹⁴. Furthermore if IPOs jump up on their first day of trading, only for these gains slowly to be dissipated in coming months and years, IPO underperformance would still be an inadequate description of events.

Potentially even more seriously (for hypothesis testing), the same experiment reveals that conditioning on prior data availability reduces the variance of monthly returns, which leads to overrejection of the null of no abnormal performance.

On the other hand, the variance of *average* abnormal returns over the test period will also change as the sample size changes due to drop outs. For example, if there is no abnormal performance as hypothesised under the null, we expect the *CAAR* to follow a random walk. If it is possible to assume that the variance of abnormal returns is constant over time, and that the number of firms in the sample is fixed, then the variance of $CAAR_T$ will be T times the variance of $CAAR_1$. The number of firms in the sample is not, however, fixed, because firms drop out. The variance of AAR_t is therefore not constant but increasing over time as the sample becomes smaller. Hence this is a reason to expect the estimation period variance to be lower than the test period variance.

The Barber and Lyon (1997) simulations involve 1000 samples of 200 event months taken from the population of all monthly returns existing on CRSP records, and investigate further the specification of t-tests using \overline{BHAR} s and CAARs. They report the specification of \overline{BHAR} s using 7 methods to adjust for normal returns and CAARs using 8 methods to

sizes less than 2000

¹⁴ It is well-known, if not necessarily completely understood, that closed end funds are issued at a premium to net asset value and then tend on average to trade down to a discount (compared to net asset value) during their first months /years of trading. This is generally treated as a different phenomenon in finance and the subject has been analysed carefully by, for example, Lee, Shleifer and Thaler (1991)

adjust for normal returns, each over 3 time periods and at 3 significance levels. That is, they report the specification of 135 different test procedures. CAARs have positive mean but \overline{BHAR} s have negative means. Both are subject to new listings bias which tends to raise the mean of the statistic. \overline{BHAR} s are far more skewed than CAARs and the skewness bias means all \overline{BHAR} s except those which are based on a control firms model of expected return are misspecified and over-reject in favour of negative abnormal performance. Rebalancing bias leads to negative \overline{BHAR} s but has an ambiguous effect on CAARs. This is because spurious negative autocorrelation among the constituents of the reference portfolio depresses the value of the \overline{BHAR} when the reference portfolio is rebalanced, whereas in the CAAR calculation such effects can apply equally to the sample firm and the reference portfolio. Rebalancing bias is more serious using the US indices which are rebalanced monthly (in the UK, FTSE Allshare is rebalanced yearly, and FTSE sector indices and FTSE100 are rebalanced quarterly)

Barber and Lyon claim the clear result that of the 135 tests examined those based on control firm models and using \overline{BHAR} statistics perform best. Indeed, the extent of misspecification of such tests is very slight. One major explanation for this superior performance is that skewness bias is largely eliminated by using a control firm approach, because control firm returns, unlike index returns, are likely to be just as skewed as sample firm returns.

All of the studies discussed here find at various points that parametric tests will often identify long-run abnormal returns when none is present. KW, in their conclusion, suggest that bootstrapping techniques will improve matters. Lyon, Barber and Tsai (1998) address the issues identified by KW and BL and advance the research significantly in two

directions. Firstly, they build their own size and book to market value sorted reference portfolios based on buy and hold returns. When reference portfolio performance is measured using buy and hold returns there can be no rebalancing bias. New listings bias is also largely eliminated because for every constituent member of every sample they rebuild a reference portfolio, and then throughout the event period they ignore all new issues¹⁵. Secondly, they answer the call of Kothari and Warner and analyse bootstrapped statistics. The key finding is that after taking such care in performance measurement and test construction they can identify several procedures which lead to well specified test statistics in random samples. But once they progress to non-random samples (by non-random samples they mean, for example, samples characterised by industry clustering, or non-random levels of pre-event performance), inference remains as treacherous as ever.

The test statistics which are well specified in random samples are the t-test based on size and book to market value matched control firms as originally advocated in Barber and Lyon (1997), and two bootstrapping methods. The first of these is described as the empirical p-value, and the second as the bootstrapped skew adjusted t statistic. Both are described below.

The empirical p-value:

- 1) measure the \overline{BHAR} of a sample against the \overline{BHAR} of the size and book to market value matched reference portfolio.
- 2) For every firm in the original sample, randomly select another firm from the same size decile and book to market value quintile, and thereby construct a pseudo-portfolio
- 3) measure the abnormal return on the pseudo portfolio

¹⁵ An IPO issued shortly before the event date may, however, be part of the reference portfolio.

- 4) repeat steps two and three 999 times
- 5) use these 1000 observations to approximate the empirical distribution of \overline{BHAR} s, and determine the appropriate critical values for hypothesis testing

The bootstrapped skew-adjusted t statistic,

- 1) take a sample of \overline{BHAR} s and calculate the following skewness adjusted t-statistic;

$$t_{sa} = \sqrt{n} \left(\frac{\overline{BHAR}_T}{\hat{\sigma}} + \frac{1}{3} \hat{\gamma} \left(\frac{\overline{BHAR}_T}{\hat{\sigma}} \right)^2 + \frac{1}{6n} \hat{\gamma} \right),$$

where $\hat{\gamma} = \frac{\sum_{i=1}^n (BHAR_{iT} - \overline{BHAR}_T)^3}{n \hat{\sigma}^3}$, an estimate of the coefficient of skewness

- 2) draw (with replacement) 1000 resamples of size $n_b = n/4$ from the original sample
- 3) for each resample calculate the bootstrap resample analogue t-statistic

$$t_{sa}^b = \sqrt{n_b} \left(\frac{\overline{BHAR}_T^b - \overline{BHAR}_T}{\hat{\sigma}^b} + \frac{1}{3} \hat{\gamma}^b \left(\frac{\overline{BHAR}_T^b - \overline{BHAR}_T}{\hat{\sigma}^b} \right)^2 + \frac{1}{6n_b} \hat{\gamma}^b \right),$$

where $\hat{\gamma}^b = \frac{\sum_{i=1}^{n_b} (BHAR_{iT}^b - \overline{BHAR}_T^b)^3}{n_b \hat{\sigma}^3}$

- 4) determine empirically the critical values for t_{sa} from the relevant percentiles of the 1000 observations of t_{sa}^b .

These two test statistics appear to perform well in random samples, and they both have significantly higher power than t-tests based on size and book-to-market-value matched control firms. LBT find, however, that for all of the statistics analysed in their study, empirical and theoretical size diverge in the face of the following complications: industry clustering, nonrandom levels of preevent returns performance, book to market value, and to a lesser extent, size. In many cases, the least misspecified statistic is the t-statistic based on size and book to market value matched control firms. The bootstrapped skew-adjusted t-statistic is also relatively reliable.

2.3.2 Calendar time techniques revisited

Mitchell and Stafford (2000) argue that the problem of cross-correlation of abnormal returns within a sample is the overriding one for long-run event studies. If abnormal returns are cross-correlated it would appear that individual *BHARs* or *CAARs* are not independent and therefore that ordinary t-testing procedures on their means are invalid. They recommend a calendar time regression of the type described in 2.2.2c above or, alternatively, an adjustment to the standard t-tests (or, presumably, to the skew-adjusted t-tests discussed above). Their adjustment involves estimating the average cross-correlation of abnormal returns as a function of the extent of calendar time overlap between firms in a sample and then applying the following formula for the standard deviation of the sample mean when the individual observations have equal variances but are not necessarily independent¹⁶:

$$\sigma_{BHAR} = \sqrt{\frac{1}{N} \sigma_i^2 + \frac{(N-1)}{N} \overline{\sigma_{i,j}}} = \sqrt{\frac{\sigma_i^2}{N} (1 + (N-1) \overline{\rho_{i,j}})}$$

Specifically, they calculate all possible pairwise correlations of monthly and annual *BHARs* between pairs of firms which are in the same sample and share the same issue month. Ie, within a particular sample, any two firms whose event took place in the same month offer 3 years of monthly abnormal returns data which can be used to calculate a correlation coefficient¹⁷. The grand average pairwise correlations for such firms are quite low, (0.017 for the sample of seasoned equity offerings for example), but still potentially important. The next step is to use this estimate to impose a covariance structure on the entire dataset: They assume zero correlation of abnormal returns for pairs of firms whose 36 months of returns data do not overlap in calendar time and extrapolate linearly as the number of months of calendar time overlap increases up to the average correlation for firms with the same issue month. Thus if the 3yr *BHARs* of two firms from the SEO sample overlap in calendar time by 18 months then the cross-correlation of abnormal returns for these two firms is estimated as 0.0085 ($=0.017 \times 18/36$). At this point, armed with an estimate of abnormal returns correlation for every possible pair of firms in the sample, they take the average correlation across the entire sample and use this number to adjust the traditional (independence assuming) t-statistic. In the case of the SEOs sample the average correlation is 0.00351, a small number but significant when combined with the large sample size in the formula above. Testing the null of no abnormal performance of SEO firms they find that taking account of cross-correlations in this way changes the t-statistic from -6.05 to -1.49.

This is a dramatic change and may indicate a serious problem in the widely favoured *BHAR* / bootstrap methodology. We have mentioned that LBT (1998) describe the measurement of long-run abnormal stock returns as “a treacherous business” but the

¹⁶ Non-independence, unlike non-normality, is a problem which grows with sample size as is clear from the formula.

measurement of cross-correlations in abnormal stock returns for firms within a sample may also be extremely difficult. MS estimate pairwise correlations between all firms whose events occurred in the same month. Their large samples ensure that there were many such relationships but for each relationship there were just 36 datapoints. The linear extrapolation technique for estimating cross correlations between firms whose event periods do not completely overlap is obviously *ad hoc* but may still serve as a useful approximation.

Mitchell and Stafford strongly advocate a calendar time approach to the measurement of long-run stock performance. Ritter and Loughran (2000) make a variety of contributions to debates this area. They don't directly take issue with the MS techniques- and indeed their earlier investigation of IPO and SEO underperformance (Ritter and Loughran (1995)) includes results from calendar time portfolio regressions – but they make a robust defence of the \overline{BHAR} methodology. One of their key points is that all calendar time techniques have low power to detect abnormal performance if it is the case that periods of high event incidence coincide with periods of high abnormal performance. Ritter (1991) and Ritter and Loughran (1995) find that underperformance of new issues which arrive in months of high issuance is more pronounced. The significance of this underperformance is understated if these high volume months receive an equal weighting in calendar time regressions.

In addition, as mentioned above, Brav and Gompers (1997) found that when IPOs in a sample very similar to that of Ritter (1991) are value-weighted the average

¹⁷ The sample sizes of MS (2000) are huge (4911 seasoned equity offerings allow 50241 pairwise correlations to be estimated between pairs of firms experiencing events in the same month)

underperformance is vastly reduced suggesting that it is the smaller firms in the sample which have delivered most of the underperformance.

With these results in mind Ritter and Loughran (2000) carry out an elaborate simulation exercise which attempts to capture the observed cycles in event incidence and the positive relationship between event incidence and extent of mispricing and the level of the stock market itself. (In the case of IPOs this is known as “hot issues markets”: new issue volume appears to increase and these issues appear to perform particularly poorly over the long-run at times when the stock market is high and sentiment is bullish). Ritter and Loughran’s simulations involve drawing a firm randomly for each month between 1973-96 and inducing artificial abnormal performance (by adding 5% to each firm’s price). In addition, if in any month a broad share price index has risen during the previous 12 months by more than 30% Ritter and Loughran randomly select another 5 firms for that month from a smaller universe which excludes the biggest 20% of firms and add 100% to the price of each firm. Thus each sample displays clustering which is related to overall market movements and has the properties that misvaluations are more prevalent in small firms than in large firms and are more extreme in high volume periods. The artificially induced abnormal performance is allowed to decay linearly month by month over 36 months so that the price in the 36th month (the last month in which a sample firm enters the calendar time portfolio) is the real market price. RL calculate abnormal return using various techniques to see whether or not the misvaluation is detected. They find that generally it is but that as expected the Fama and French calendar time procedure has the least power to detect it: over 1000 samples of such overpriced firms the average implied underperformance of the Fama and French model is the lowest and furthermore the intercepts for the value weighted application of this model have the highest standard errors making a statistically significant

result less likely. The most accurate measure of abnormal performance is the \overline{BHAR} measured relative to size adjusted control portfolios.

To summarise the main points raised in this section, LR appear to have shown that the Fama and French calendar time portfolio regression techniques have low power to detect abnormal performance. On the other hand this technique is certainly immune to the potential problem of cross correlation of abnormal returns within a sample. If cross correlation is present then hypothesis tests based on the \overline{BHAR} and bootstrapped t-statistics may have inaccurate size. The real question seems to be, do these cross-correlations actually exist?

The big selling-point for calendar time techniques in long-run event studies is that if cross correlations in abnormal returns of firms within a sample exist then they will be properly accounted for by a calendar time procedure. On one level it is easy to interpret the presence of industry clustering, calendar time clustering etc in samples as evidence that firms' have particular reasons for undertaking particular corporate events at particular times: they sell equity when they perceive the value of their equity to be high, for example. If it is the case that at certain times lots of similar firms sell equity at prices which are irrationally high and if these irrationally high prices are unsustainable in the long term then negative abnormal returns will ensue and individual firms' abnormal returns will be correlated. This is true but such an interpretation raises further difficult questions: how do prices become irrationally high, and why is there investor demand for new issues when prices are irrationally high? Cross-correlation of abnormal returns implies irrational pricing, and calendar time techniques rely on cross-correlation of abnormal returns for their advantage. Even so, as RL (2000) show, calendar time techniques lead to hypothesis tests with low power.

Table one presents a summary of the findings in this chapter.

Table 1. Summary of Findings.

Panel A: event time techniques (CAARs and \overline{BHAR} s)

<i>Cause of misspecification</i>	<i>Does the problem apply to CAARs?</i>	<i>Does the problem apply to \overline{BHAR}s?</i>	<i>What is the consequence? Is the problem surmountable?</i>
Skewness bias	Yes	Yes (even more than in CAARs)	Test statistics biased towards negative abnormal performance. Any control firm approach eliminates skewness bias (as does skew adjusted bootstrapped t statistic)
New listings bias	Yes	Yes	CAARS and \overline{BHAR} s biased upwards. The hugely labour intensive LBT process of reference portfolio construction largely eliminates the bias. Again, new listings bias is less serious using control firms approach
Rebalancing bias	Yes (bigger problem in CAARs than in BHARs)	Yes, slightly, and not at all under a control firms approach	\overline{BHAR} s biased downwards, ambiguous effect on CAARs. Bias can be eliminated by the labour intensive LBT process of reference portfolio construction
Non-random samples (all types but especially industry clustering and nonrandom levels of preevent returns performance and book to market value ratio)	Yes	Yes	Sample specific. LBT don't detect any reliable patterns but empirical size of hypothesis tests likely to be larger than theoretical size.
Calendar time clustering	Yes	Yes	Downward biased variance estimates leading to over-rejection of nulls. Bootstrapped statistics developed in LBT may improve test specification (still contentious issue)

Panel B: calendar time techniques (CCTAR or Fama French calendar time portfolio regressions)

Advantages	Event time techniques may generate t-tests based on faulty standard errors which don't take account of cross-correlations in abnormal returns within a sample. Only calendar time techniques can properly account for these correlations (if they exist). In a nutshell, accurate size.
Disadvantages	Low power

2.4 Conclusion

Fama (1991) states that event studies offer the cleanest evidence on market efficiency, as they are the least encumbered by the joint hypothesis problem. The issues analysed in this chapter suggest strongly, however, that the remark only applies to studies which investigate a short post-event window. Studies investigating the long-run performance of IPOs, stock splits etc, are as dogged by the joint hypothesis problem as any other tests of market efficiency. In addition there are several other problems which arise in long-term studies and which can lead to badly biased test statistics with low power or alternatively a tendency to over-reject the null.

As a result of the research discussed in this chapter, much more is now known of the performance of the various test statistics which have been used in long-run event studies, and it is possible to draw some conclusions and to make some pragmatic recommendations: among event time techniques \overline{BHAR} s appear more susceptible to common specification problems but may be preferable to CAARs if steps are taken to alleviate known misspecification since they measure better the concept of interest (ie *long-run* abnormal returns); control firm techniques or the elaborate reference portfolio construction techniques of LBT are better than any procedure relying on “ready-made” index data; the bootstrapped skewness adjusted t-test or empirical p-value techniques described by LBT should be employed. Even these statistics, however, will not be completely robust to (the likely) nonrandomness of a sample. Both value weighted and equal weighted \overline{BHAR} s should be presented. If abnormal performance is identified then a calendar time technique such as the Fama and French calendar time portfolio regression should also be included in the analysis. If this latter procedure does not identify abnormal performance which is indicated by \overline{BHAR} analysis the results of the investigation are inconclusive as the issue of cross-correlation of abnormal returns within a sample remains unsettled.

Chapter 3. Long–Run Performance Analysis of a New Sample of UK IPOs

3.1 Introduction

In the last chapter we discussed in some detail the econometric issues which complicate a long run event study. Here we refocus on IPO performance and carry out an empirical investigation of the abnormal performance of a new sample of IPOs which have been issued in UK. We incorporate in the methodology as many as possible of the recommendations noted in the previous chapter.

There are three features of the performance of initial public offerings (IPOs) which have been highlighted. They are that initial returns on IPOs are, on average, abnormally high; that the magnitude of this abnormally high initial return is variable, indeed cyclical; and that in the longer run, IPOs yield significantly *negative* abnormal returns. The first two propositions are more strongly supported by empirical evidence than is the third. This chapter contributes to the research on the three issues, but focuses on the third, long-run abnormal performance. The principal conclusion is that the IPOs in this new dataset do not display general long-run abnormal performance which is statistically significant.

Broadly, there are three ways in which the chapter builds on or sheds further light on the results of existing research in this area. Firstly, the new dataset examined here covers IPOs issued in the UK between 1990Q2 and 1995Q1, and is as recent as possible. This is interesting because other efficient markets anomalies such as the size effect and certain patterns of serial correlations in stock returns have tended to get weaker over time, and it

may be that a parallel trend is occurring in long-run IPO performance¹. Secondly, since there have recently been highlighted various problems with both the measurement of abnormal returns and the specification of tests for non-zero abnormal returns, (as discussed in the previous chapter), a bootstrapped skew-adjusted test statistic has been employed. Thirdly, results are presented on abnormal returns using a procedure which controls for industry sector effects on expected return. Industry sector is not, admittedly, a risk factor which is frequently suggested in portfolio theory except to the extent that firms within an industry have identical exposure to systematic risk. However the analysis throws up some curious results which may be of interest to the proponents of behavioral finance and “noise trader” theories in particular: the behavior of information technology related IPOs, (quite plausibly those which attracted the most attention from “noise traders” during the period) is quite different from the behavior of IPOs in any other industry sector. Information technology related IPOs experience initial returns far in excess of IPOs in any other industry sector, and then drastically underperformed the rest of their own industry sector during the subsequent three years. Both of these findings *are* statistically significant.

The central hypothesis under investigation is that IPOs underperform in the long-run: on average, they yield significant negative abnormal returns during their first few years of trading. Evidence of this anomaly has been documented, for example, by Ritter (1991) and Ritter and Loughran (1993) using US data, by Levis (1993, 1995), and Espenlaub *et al* (1997) using data from the UK, and by Lee Taylor and Waltor (1996) on the basis of data from Australia. It’s an intriguing hypothesis which appears to contradict fundamental efficient markets tenets. Confirmation would raise a difficult question for economists : if

¹ A good example is the “January effect”, (the observation that stock returns have tended to be higher in January than in other months), which has been analysed by, among others, Keim, (1989).

IPOs consistently and significantly underperform during their first few years of trading, why does anyone ever buy them?

The matter of long-run performance is germane to research on IPO performance more broadly. *Positive* and highly significant *initial* abnormal returns for IPOs (meaning returns over the first day or two of trading), are an empirical phenomenon almost universally accepted, and confirmed in the present dataset. There has grown a large theoretical literature which seeks to explain this phenomenon². Frequently the premise is that IPOs are underpriced by firm owners and/or their advisers, and the theories seek to explain this underpricing. If, however, it is confirmed that IPOs perform poorly during the early years of trading, then this basic premise in the analysis of initial returns must be faulty. A satisfactory theory of IPO performance would not explain why IPOs are underpriced at issue, but why the open market price jumps up irrationally when the stock starts trading, only for these initial gains (and more) to be dissipated slowly over subsequent years. This would appear to be a far more perplexing puzzle. To summarize this point, existing evidence of short-run overperformance and long run underperformance of IPOs suggests that theoretical research is incomplete or misguided if it seeks *only* to explain IPO *underpricing*. If the results on long-run performance presented in this paper gain credence, however, IPO underpricing *is* in fact the only empirical puzzle.

Existing research on long term returns has tended to follow the IPOs for 3 or 5 years after issuance. No particular length of post-event window is suggested by theory. In order to permit as recent a dataset as possible a post event window of 3 years has been studied. As a result the dataset at hand overlaps very little with IPO datasets used in existing research.

² See Ibbotson and Ritter (1997) for a very comprehensive survey

It is also the case that since measurement error in the calculation of abnormal returns is inevitable, and since this error must increase with the length of the post-event window, tests which are directed at longer horizon returns are more susceptible to the econometric problems discussed in the previous chapter – hence the need to take account of these problems.

Section 3.2 describes the data. Section 3.3 discusses methodology and test statistics. Section 3.4 presents results. Section 3.5 concludes.

3.2 Data

The starting point for data collection was KPMG's quarterly publication "New Issue Statistics", which reports all new issues of stock to markets organized by the International Stock Exchange in London. After excluding rights issues, seasoned equity offers, investment trust issues and the government's privatisation issues³, returns information was collected on the remaining issues using Datastream. Datastream was also the source of the FTSE Allshare index and the FTSE sector indices data. Ideally, all IPOs would be included, (including those on the Unlisted Securities Market), but Datastream does not carry researchable historic price information for all listed firms. Data was collected successfully for 232 out of 288 IPOs fitting the criteria set out above; the equity raised in the 232 included offerings was £10.137bn which is 91.36% of the total amount raised in

³ Espenlaub, Gregory and Tonks (1997) exclude privatisation issues from their IPO dataset, but Levis (1993) does not. It is arguable that the motivations for the decision to sell, and indeed the pricing decision itself, may be different in the case of government privatisations. For example, the government may deliberately offer shares cheaply "to leave a good taste in investors' mouths", if it cares more about encouraging new shareholders than does a private firm. As a practical matter, the main suspicion must be that privatisation issues are offered cheaply in order to help the government's popularity with investors and as such we suspect that government IPO's long-run returns may have a positive bias. Since we have in mind a one tailed alternative hypothesis that IPOs *underperform* on average, the exclusion of such issues from the dataset should allow

the 288 offerings.

The root of the missing data problem is that many IPOs are extremely small. Frequently floated by way of placement to a small number of specialist institutional investors, such issues can be highly illiquid after flotation and, in fact, very rarely traded. When such a stock doesn't trade in the market for days on end, a daily closing price series, even if available, may not be particularly insightful.

Datastream generates a total returns index which is ideal for the calculation of \overline{BHAR} s and CAARs. The index is available at daily frequency. It incorporates all dividends on the appropriate dates and is adjusted for scrip issues, rights issues and other recapitalisations.

Table 3.1 presents summary statistics on the dataset. IPO issuance clearly varies with the economic cycle: issuance is low in the early years of the sample, a period in which the UK economy was close to, or in, recession. Stock market valuations at that time were relatively depressed. Table 3.2 contains the same data, sorted by market sector rather than year of flotation. Compared with earlier datasets used in existing research, Table 3.2 illustrates an increase in information technology related issues and a decline in the importance of extractive industries.

Table 3.1: The Data

year	number of IPOs	total funds raised (£m)	average funds raised per IPO (£m)	total market value on flotation (£m)
1990-91	9	32.372	3.60	1751.47
1991-92	13	748.585	57.58	1648.23
1992-93	28	1701.795	60.78	4219.36
1993-94	98	3808.191	38.86	7871.27
1994-95	84	3846.048	45.79	11014.27
1990-95	232	10136.991	43.69	26504.60

Table presents data on all IPOs (excluding investment trusts) which were issued in London during the period beginning 1990 Q2 and ending 1995 Q1 and for which Datastream presents researchable price history.

Table 3.2: The Data by Industry

Market Sector	number of IPOs	total funds raised (£m)	Average funds raised per IPO (£m)	total market value on flotation (£m)
healthcare and pharmaceuticals	23	576.689	25.073	1475.768
food production, forestry and paper, packaging, engineering, autos, chemicals, construction and building materials, diversified industrials	43	2421.862	56.322	5010.710
oil, gas, mining	10	160.165	16.017	391.630
transport and distribution	27	735.912	27.256	1453.170
beverages, restaurant, leisure, media	33	1743.448	52.832	6398.440
banks, insurance, real estate and speciality financial	34	1327.924	39.057	3688.120
electronics, infotech hardware, computer software services, support services, telecommunications	40	1573.621	39.341	5254.283
retailing, stores, household goods	22	1597.370	72.608	2832.480
All	232	10136.991	43.694	26504.601

This table contains the same data as Table 3.1 above, but the data are sorted by industry sector rather than by financial year of issue

3.3 Methodology

3.3.1 Abnormal Returns

As discussed in detail in the previous chapter, in order to measure abnormal return, it is first necessary to have some notion of normal return. This is not an important matter in the analysis of initial returns, since the post event window is just a day, but it is very important in the analysis of long-run returns.

In this chapter, the analysis of *initial* returns uses raw returns data, i.e. returns which don't take any account of what a normal return, (expected return), might be. As illustrated in Section 4 below, IPO raw returns on the first day of trading are of such a magnitude and variance that differences in the underlying model of expected return or "normal return" are of second order.

For any analysis of long-run returns, however, a model of expected return is most certainly required. The model which has been used most frequently in existing research on IPOs is the market-adjusted returns model. This model measures abnormal return in a particularly straightforward way:

$$(1) \quad ar_{it} = r_{it} - r_{mt}$$

Abnormal return is the raw return on the IPO minus the return on the market during a particular period. Usually, the return on the market will be measured by a broad share price index, such as the FTSE Allshare Index.

Clearly, this model has the advantage of simplicity, but just as clearly it is not an accurate representation of any conventional portfolio theory. Any such theory of expected asset returns would predict more cross-sectional variation in expected return. The market adjusted returns model implies that the expected return of any asset, or any portfolio for that matter, is the same. In the context of the traditional CAPM for example, this implies the restriction that the beta coefficient for any and all assets is 1. The unrestricted CAPM suggests that an asset's expected return depends positively on its covariance with returns on the market portfolio. Some writers have calculated an estimate of the CAPM beta using ex post price data⁴, but this approach suffers from a drawback in that it imposes an extra data requirement (of at least two years) for parameter estimation. Furthermore, deviations from the CAPM have been well documented, and appear to be frequent. Applications of the CAPM in IPO research have suggested that beta coefficients may not be stable over time⁵. Nor has there emerged any other equilibrium model of asset prices which performs well enough empirically to gain the support of a majority. Of course any model is an imperfect description of reality so an amount of measurement error is inevitable. All models are wrong but there is evidence to suggest that the market adjusted returns model is often the model under which the size and power of subsequent hypothesis tests is least misspecified⁶. From this pragmatic point of view, the primitive market-adjusted model of abnormal returns may be as good as a more complicated procedure.

There is, though, at least one caveat: a sample of IPOs typically includes a very wide range of firm size. It is unlikely to reflect accurately the size composition of any particular market index, so calculations of abnormal returns may not be valid if firm size is a

⁴ See Espenlaub, Gregory, Tonks(1997) for example.

⁵ Ibbotson (1975), Levi (1995)

⁶ As discussed in Chapter 2 this was one of the conclusions of Brown and Warner (1980), supported more recently by Barber and Lyon (1997)

determinant of return⁷. In the present sample more than half of the IPOs are capitalized at £50m or less at the initial offering price, while at the other end of the spectrum the average initial market capitalization over the largest 5 IPOs is over £2bn per firm. Thus the present sample has a higher proportion of small stocks than the FTSE Allshare.

Even if size doesn't matter there are further problems with the calculation of statistics based on index adjusted abnormal returns. These problems are taken up below.

The alternative approach to the measurement of abnormal returns does not refer to a stock market index but instead involves selecting matching firms for each sample firm and comparing sample firm performance to the average performance of the matching firms. The matching firms should control for as many potential determinants of expected returns as possible in order that the abnormal return figures isolate and capture only the effects on the firm of being an IPO. Although strictly we do not apply this technique in the present study, the extension in which we control for industry sector in the measurement of abnormal return is in the same spirit.

3.3.2 Test Statistic

The test statistic which has proved most popular in existing literature is the cumulative average abnormal return (CAAR) but here, for reasons discussed in the previous chapter, the results are presented using buy-and-hold abnormal returns (*BHARs*).

⁷ Relative returns on small stocks and large stocks have certainly varied widely in the past. For a long time there was evidence that on average small stocks yielded significantly higher returns than large stocks. This has been known as the size effect, but it is an anomaly which is undetectable in recent data. In fact small firms have on average yielded lower returns than large firms during most years in the last decade in both the US and the UK.

The T period buy and hold abnormal return for a given sample firm is

$$(2) \quad BHAR_{iT} = \prod_{t=1}^T (1 + r_{it}) - \prod_{t=1}^T (1 + r_{mt})$$

The simple average over an entire sample is

$$(3) \quad \overline{BHAR}_T = \frac{1}{n} \sum_{i=1}^n BHAR_{iT}$$

and a conventional test statistic would be

$$(4) \quad t = \overline{BHAR}_T / \sigma(BHAR_{iT}) * n^{1/2}$$

3.3 Potential biases

As discussed in detail in the previous chapter, Brown and Warner (1980) presented a step by step good practice guide for event studies. They highlighted several reasons to suspect test statistic misspecification in many long-run event studies. The issues raised have proved to be only the tip of an iceberg. More has been discovered recently by Kothari and Warner (1997), Barber and Lyon, (1997), and Lyon Barber and Tsai (1998) among others. Tests of long-run abnormal security performance, whether based on the CAAR or the \overline{BHAR} , are likely to be subject to a skewness bias, a rebalancing bias, and possibly a new listings bias. There are steps which can be taken and which may alleviate or occasionally remove these problems, but even if these steps are followed diligently the standard t -statistics will only be well specified in truly random samples, which may well not prevail in most event studies. Industry clustering, calendar time clustering, nonrandom levels of

pre-event returns performance and book to market value have all been shown potentially to affect adversely the size and power of such tests.

To summarize the conclusions of this research, rebalancing bias suggests that \overline{BHAR} s will be negative on average, and skewness bias suggests that the true size of standard t-tests of the zero abnormal returns null against the one-sided alternative of negative abnormal returns is likely to be higher than the theoretical size of such tests.

In order to confront these problems as much as possible the test statistics presented in this chapter are a skew adjusted and bootstrapped version of the traditional t statistic. The skew adjusted t statistic developed by Johnson (1978) can be written as

$$(5) \quad t_{sa} = t + \hat{\gamma} \left(\frac{1}{6n} + \frac{\overline{BHAR}_T^2}{3\hat{\sigma}^2} \right) \sqrt{n}$$

where t and $\hat{\sigma}$ are the traditional t statistic and sample standard deviation from equation

(4), and $\hat{\gamma}$ is the sample coefficient of skewness given by $\hat{\gamma} = \frac{1}{\hat{\sigma}^3} \cdot \frac{1}{n} \sum_i (BHAR_{iT} - \overline{BHAR}_T)^3$.

Sutton (1993) demonstrates that there are various computer intensive techniques based on bootstrap resampling which can improve the performance of t_{sa} yet further⁸. The form of statistic presented here is a variation of the normal approximation method and is one of the statistics recommended by Sutton (1993) for testing the mean of an asymmetric distribution⁹. It proceeds as follows:

⁸ Essentially, bootstrap re-sampling involves treating the original sample as a population, and re-sampling from it with replacement.

⁹ Sutton's bootstrapped skew adjusted t-statistic has been employed here in preference to the one described by LBT (2000) and outlined in Section 2.3.1 above. LBT refer to Sutton's paper but do not explain why their own version of the bootstrapped skew-adjusted statistic differs from that of Sutton. Furthermore they note without explanation (footnote 8) that their statistic is well specified

- i) from the original sample of 232 IPOs take 1000 bootstrap resamples each of size 232 and for each of them calculate the (skew adjusted) t statistic as in (5) above.
- ii) Calculate the standard deviation s_t of these 1000 t -statistics
- iii) Calculate the ratio t_{sa} / s_t using the (skew-adjusted) t -statistic from the original sample, and compare to the critical values of the standard normal distribution.

This procedure is only justified when the null hypothesis sampling distribution of the test statistic is approximately standard normal. Hence the skew-adjustment is crucial: this form of bootstrapping procedure will generate badly misleading results if applied to the conventional t -statistic when the underlying observations come from a distribution which is skewed (because in such cases the sampling distribution of t is itself skewed in the opposite direction).

Lyon, Barber and Tsai, (1998) recommend skew-adjustment and bootstrapping when testing long-run security returns but these authors are quick to point out on the basis of their research that elements of non-randomness in the sample can still lead to inaccuracies in the size of the test. Indeed they state in their conclusion that “Our central message is that the analysis of long-run abnormal [security] returns remains treacherous.” Unfortunately, a sample of IPOs is unquestionably nonrandom: private firms’ decisions to go public are likely to be related to economy wide factors such as the overall level of the stock market and credit conditions as well as to factors such as product market conditions and stock market valuations of particular industry sectors. In this paper, these problems are confronted pragmatically: as many results as possible are presented, using different

when the resample size is $n/4$ or $n/2$ but not when it is n . This is disconcerting even though I don’t find evidence of misspecification in a primitive Monte Carlo simulation (not presented in this thesis)

benchmarks for expected return and different test statistics.

3.4 Results

Results on the abnormal returns of the IPOs in the present dataset are reported in two stages: initial returns and long-term returns are examined separately. The analysis is split in this way because existing research suggests that the initial return is abnormally positive yet the long-term return is abnormally negative. If studied together, the one will mask the effect of the other.

The closing price on the first trading day is arguably the most appropriate place to begin measurement of long-run performance of IPOs for a second reason: in many cases it is not possible for non-specialist investors to buy stock at the initial offer price so the abnormal return measured from the closing price of the first day's trading is an achievable return whereas in some cases the return measured from the offering price may not be.

3.4.1 Initial Returns and the "Hot Markets Phenomenon"

Tables 3.3 and 3.4 summarize the results on initial returns of the IPOs in the present dataset. Initial return is defined as the return from buying shares at the initial offer price and selling them at the closing price on the first day of trading. An overall average 1st day return of 8.70% is in line with previous studies. Ibbotson and Ritter (1997) survey they evidence from many countries and report initial returns on new issues to stockmarkets in developed economies running from 4.2% in France up to 15.3% in USA. In the famous study by Ritter (1991), the average raw first day return on the sample of 1526 IPOs was

when the resample size is n .

14.32%.

As stated the returns reported here are raw - there has been no adjustment for expected return. The average daily return on the FTSE Allshare index over the sample period was 0.045%¹⁰, far lower than the initial returns on IPOs, so any adjustment according to a model of expected returns would make virtually no difference. The median initial return and value-weighted average returns yield further insights. The median return is lower than the (equal weighted) average return suggesting that the distribution of initial returns is skewed to the right¹¹, as is widely documented in existing research. This is easily discernible from Figure 3.1, the histogram of initial returns. The value-weighted average returns are yet lower, though still substantial, indicating that on average it is the smaller stocks which have the very high initial returns. (In the value weighted calculations the smaller stocks receive lower weights). Over the entire sample, the equal-weighted average initial return exceeds the value-weighted average by a factor of 1.75, which suggests that size is an important determinant of initial return. Viewed loosely, and from a different perspective, we may note from these findings that the strategy of placing a level stake in all IPOs in the sample would have beaten the strategy of spending the same sum of money on an equal proportion of the share capital of each of the floated firms, (if the positions were all sold at the first day's closing prices).

¹⁰ Source: Datastream

¹¹ This finding was recognized and discussed as early as 1975, (Ibbotson, (1975))

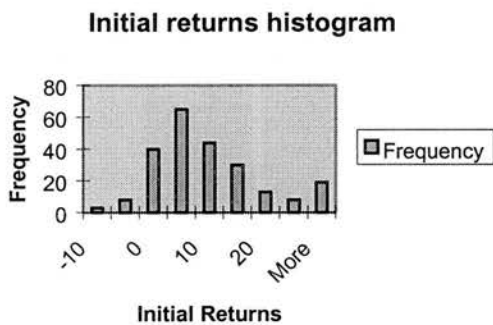
Table 3.3: Initial returns

year	no of new issues	(equal-weighted) average initial return (%)	median initial return(%)	(value-weighted) average initial return (%)	12 month BAHR on FTSE Allshare (%)
1990-91	9	7.40	7	3.35	12.46
1991-92	13	2.77	2	0.43	3.31
1992-93	28	14.22	10	6.15	26.18
1993-94	98	11.17	8.5	6.22	15.16
1994-95	84	5.03	2.45	4.60	2.79
1990-95	232	8.70	5.1	4.99	73.531*

Table shows average initial returns on the sample of 232 IPOs issued in London between 1990 Q2 and 1995Q1. Initial return is defined as the return from buying a stock at the initial offer price and then selling at the closing price on the first trading day. Data are presented raw (without any adjustment for expected return).

*73.53165% is the 60 month BAHR on the FTSE Allshare, rather than an sum of the 5 different 12 month BAHRs above.

Figure 3.1



Histogram of initial returns data from Table 3 above. Initial returns are quite clearly skewed to the right: the skew statistic is 4.5314.¹²

One may argue that value weighted averages bring the appropriate perspective to any

anomalies which are uncovered: it is worth restating that the small firms in a sample of IPOs are really very small indeed, and to use patterns in their market performance to draw inferences about the efficiency of the entire stock market would be rather rash. A second consideration is that asset pricing models such as CAPM or the Fama and French 3 factor model perform less well in explaining the returns behaviour of small stocks¹³. When we come to adjust the raw returns for expected return in the long-term analysis, bad model problems will inevitably interfere with test statistics, and to the extent that these problems are more severe in the case of small stocks, it may be reasonable to be more interested in the value weighted results rather than the equal weighted results.

The “hot markets phenomenon” refers to the observation made by Ritter (1991) and others, that the size of average initial returns on IPOs appears to vary over time. Indeed it appears to vary with the economic cycle: average initial returns are high when the economy is growing and the stock market bullish. The present sample does not span a long enough time period to permit a rigorous examination of the “hot markets phenomenon” but it is clear that during 1992-94, years when the stock market was particularly strong, IPO issuance was heavy and average initial returns were high¹⁴.

Similar patterns emerge in Table 3.4 when the initial returns are sorted by industry sector: all market sectors delivered positive initial returns. The extreme outlier is the information technology and telecommunications sector, mainly composed of fairly small offerings,

¹² This statistic is calculated as follows: $Skew = \frac{n}{(n-1)(n-2)} \sum_{i=1}^n \left[\frac{IR_i - \overline{IR}}{s} \right]^3$ where IR stands for (gross) initial return, and s is the sample standard deviation.

¹³ See, for example, Brav and Gompers (1997)

¹⁴ The 12 month BAHHR of 12.46% for the FTSE Allshare Index in 1990-91 is slightly misleading: 12.46% is a healthy annual return but the period should not be thought of as a bull market period: the economy was in recession, and business and stock market sentiment was gloomy - the relatively high BAHHR can be connected with the observation that the market fell very sharply in the last weeks of 1990Q1, (this dataset begins on the

which delivered an average initial return of 17.9%. This, of course, is a quite spectacular average one-day return. As in the data for the whole sample we see that in every market sector the average initial return is higher than the median initial return, indicating the presence of positive skewness. In all market sectors except for transport and distribution, we find that the equal weighted average initial return exceeds the value weighted average initial return, confirming that in most sectors the positive initial returns are more exaggerated in the case of small firms.

Table 3.4: initial returns by market sector

Market Sector	number of IPOs	(equal-weighted) average initial return(%)	median initial return(%)	(value-weighted) average initial return(%)
healthcare and pharmaceuticals	23	6.609	3.6	3.725
food production, forestry and paper, packaging, engineering, autos, chemicals, construction and building materials, diversified industrials	43	8.142	5.9	6.037
oil, gas, mining	10	4.140	3.95	-0.586
transport and distribution	27	6.974	6.3	8.675
beverages, restaurant, leisure, media	33	7.191	4.8	5.909
	34	5.035	2.1	1.575
banks, insurance, real estate and speciality financial				
electronics, infotech hardware, computer software services, support services, telecommunications	40	17.938	7.75	6.057
retailing, stores, household goods	22	7.268	6.25	3.029
All	232	8.697	5.1	4.986

Table 4 contains initial returns data on the 232 sample firms sorted by industry sector rather than by year of issue as in Table 3 above. Initial return is defined as the return from buying at the initial offer price and selling on at the closing price on the first day's trading.

3.4.2 Long-run Abnormal Returns

15 of the 232 firms in the sample did not survive their first three years of trading. For these 15 firms the buy and hold abnormal returns from the end of the first day's trading until the day of delisting have been measured. These abnormal returns have been used

first day of 1990Q2).

with an equal weight in the calculation of average abnormal return. The implicit assumption is that when an IPO is delisted the investor is able to switch out of the IPO at the last day's trading price, and into the stock market index.¹⁵

Table 5 shows results on long run returns for the entire sample and for subsamples of each of the financial years 1990 to 1994.

Table 3.5: Long-Run BAHRs

year	no of new issues	average 3yr buy-and-hold return on IPO(%)	average 3yr buy-and-hold return on FTAllshare(%)	3yr BHAR(%)	skew-adjusted bootstrapped t statistic
1990-91	9	98.081	47.328	50.75	1.19
1991-92	13	55.228	42.716	12.51	0.54
1992-93	28	32.505	53.22	-20.72	-0.99
1993-94	98	34.733	44.361	-9.63	-0.96
1994-95	84	79.294	70.22	9.08	0.51
1990-95	232	54.205	55.113	-0.909	-0.083

Table shows 36 month average buy and hold abnormal returns for the sample of 232 IPOs issued in London between 1990Q2 and 195 Q1. Critical values for the skew-adjusted bootstrapped t-statistic come from the standard normal distribution, so if the test is H_0 : no abnormal performance vs H_1 : negative abnormal performance (one sided) then $t_{sa} < -1.645$ is significant at 5% (none of the results are significantly different from zero at conventional levels of significance)

Whereas results on initial abnormal returns in Tables 3 and 4 above conform closely with results of earlier studies, the present results on long-run abnormal returns clearly do not. Average abnormal 3 year BAHRs vary quite wildly during the sample period, but the aggregate average abnormal performance, roughly -1%, is remarkably close to zero. In other words the IPOs in the sample tended to provide roughly the same 3 year holding

¹⁵ This seems perfectly reasonable in the cases where the delisting is due to takeover or merger etc, but unrealistic when the delisting is due to suspension pending bankruptcy. I have been unable to ascertain exactly the circumstances of all delistings in the sample. When I recalculate results assigning -100% raw return to the 5 firms I believe delisted pending bankruptcy, the average abnormal 3yr BAHr for the entire sample is slightly lower at -1.02% (the result is not much changed because these firms already display very low raw returns)

period return as the stock index over the relevant period. These returns were calculated from the closing price on the first day of trading – this means, since the average initial (raw) return was 8.70%, that if an investor had been able to buy each IPO at the *offer* price rather than the first trading day's closing price, the IPOs in the sample would have proved superior investments.

Brav and Gompers (1997) have suggested that much of the underperformance identified in earlier IPO research disappears when BAHRs are value-weighted. In fact, as illustrated clearly in Table 6, the underperformance of the present sample of IPOs is far stronger if the BHARs are value-weighted. This clearly suggests, again, that there is some kind of size effect in the data. As in the results on initial performance, but in contrast to the Brav and Gompers (1997) results on long term abnormal performance, the smaller firms in the sample have performed better than the larger firms. Of course the data used by Brav and Gompers spanned an earlier time period, and used data from different stock markets. In fact the value weighted average in the present sample is strongly affected by the presence of 2 outliers: Telewest Communications and Waste Management Group. These two firms, capitalized on flotation at £1501m and £2193m respectively, delivered 36 month *BHARs* of -131.9% and -112.3% respectively. If they are both dropped from the sample the value weighted abnormal BAHAR rises to -3.8%, which is smaller in absolute terms than the (positive) value weighted average initial return.



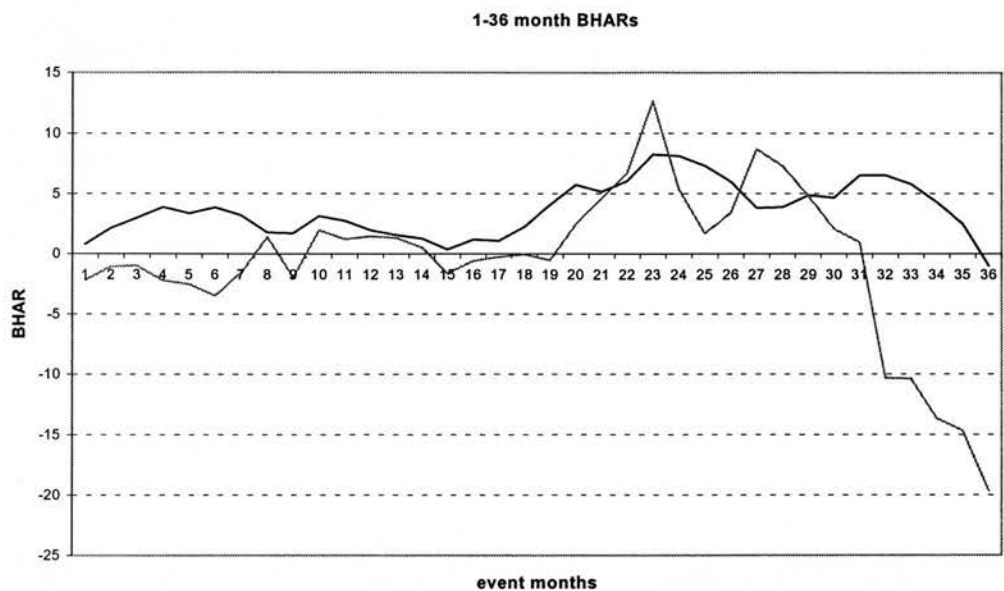
Table 3.6: long-run value weighted BHARs

Year	Number of IPOs	3-yr value-weighted BAHR	skew-adjusted bootstrapped statistic	t-
1990-91	9	-9.86294	-0.26	
1991-92	13	-7.47225	-0.53	
1992-93	28	-68.9151	-1.74*	
1993-94	98	-0.55423	-0.03	
1994-95	84	-18.7686	-1.09	
1990-95	232	-19.9504	-1.92*	

Table shows 36 month value weighted average buy and hold abnormal returns for the sample of 232 IPOs issued in London between 1990Q2 and 195 Q1. The weights in the average correspond to market capitalization at the flotation price. Critical values for the skew-adjusted bootstrapped t-statistic come from the standard normal distribution, (asterisk indicates that a value weighted BHAR is significantly below zero in a one-sided hypothesis test at 5%)

It is also of interest to look at \overline{BHAR} s at periods other than 36 months. Figure 3.2 shows value weighted and equal weighted \overline{BHAR} s measured at the end of each of the first 36 event months. The figure suggests that if one had chosen to be interested in almost any post-event window shorter than 36 months, one would have concluded that the long run abnormal performance of the present sample of IPOs was in fact positive.

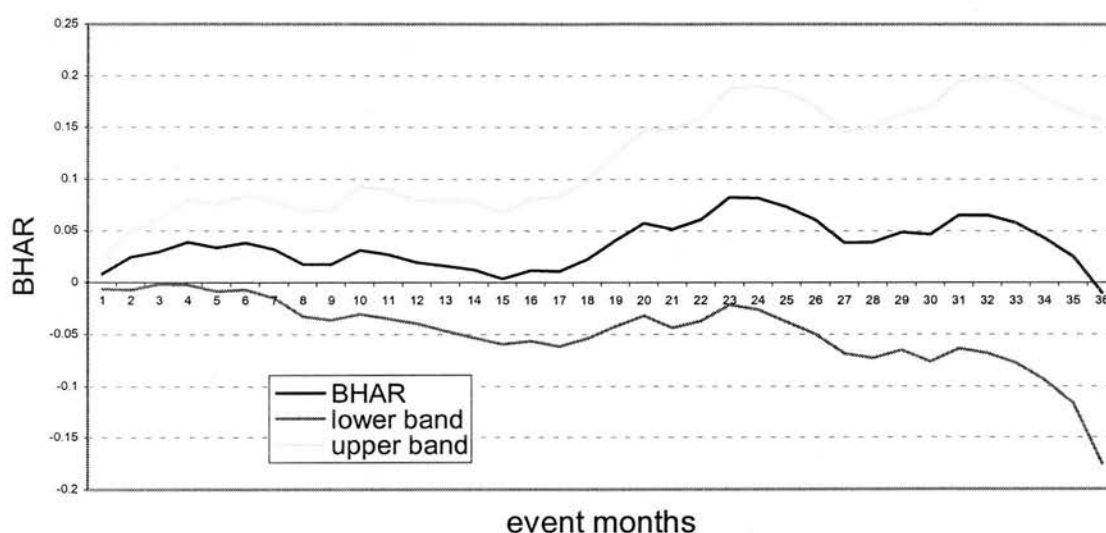
Figure 3.2: 1-36 month equal weighted and value weighted BHARS



The dark line shows the benchmark equal weighted BHAR for the full sample of 232 IPOs issued in London between 1990 Q2 and 1995 Q1 at each of the first 36 months after issue. The light line shows the value-weighted BHAR for the same sample over the same period (sample firms are weighted by market capitalization at the flotation price)

Figure 3.3 shows the equal weighted \overline{BHAR} with confidence bands. Zero is contained in the confidence interval at each of the 36 data points, suggesting little basis on which to claim abnormal performance.

Figure 3.3: (equal weighted BHARs with confidence intervals)



Dark line in the middle replicates dark line in Figure 3.2 above, plotting the buy and hold abnormal return (BHAR) for the full sample of 232 IPOs after each of the first 36 months of trading.

As in the analysis of initial returns the sample can be separated by industry sector. The third column in Table 3.7 rearranges the data of Table 3.5 above, sorting by industry sector instead of year of issue. It appears that the IPOs in the beverages, restaurants, leisure and media sector, and in the retailer and household goods sector, delivered average buy and hold returns substantially higher than the broad FTSE Allshare index. Even this abnormal performance of 44.9% and 38.8% over 36 months for these two sectors respectively, however, is not statistically significant¹⁶. Likewise, the 43 firms in the food production, forestry and paper,... etc sector delivered negative abnormal performance of 39% over three years when measured against the FTSE Allshare index, but this too is not quite statistically significant. The underperformance of the electronics, IT and telecommunications sector is also of note since the period of study included the start of a very strong rally in the value of the so-called “TMT stocks” (technology, media and telecommunications). Admittedly, the sector classification employed here does not match

¹⁶ When the sample is divided in this way, the power of the t-tests becomes low due to the small

the “TMT” sector perfectly and the abnormal performance of the IPOs in the present subsample turns positive again when the strong initial return of 17.94% (Table 3.4 above) is included but the total abnormal performance measured from the issue price is not significantly positive and this is something of a surprise.

More light is shed on this issue in the last two columns of Table 3.7 which contain the results of another experiment designed to examine the effects of controlling for market sector in the measurement of abnormal return. BHARs are recalculated for each firm in the sample using the market adjusted returns model as before, but the return on the FTSE Allshare index is replaced by the return on the appropriate FTSE sector index. This procedure is justified if industry sector is a determinant of expected return, or more plausibly, if industry return is able to proxy for expected return. This is not a common idea in portfolio theory but as stated earlier the market-adjusted returns model is clearly rather a rigid model of expected returns. Employing the model with reference to the FTSE sector indices allows for cross-sectional variation in expected returns in an intuitive way. It also gets closer to the control firms approach for measuring abnormal performance in event studies. (The idea behind the control firms approach is to estimate abnormal return by taking the difference between the sample firm return and the return on a matching firm which controls for as many firm characteristics as possible.)

Table 3.7: BHARs built using the market index, and the sector index, categorized by market sector

market sector	Number of IPOs	3yr market index BHAR (%)	Skew adjusted bootstrapped t-statistic	3yr index (%)	sector BHAR	Skew adjusted bootstrapped t-statistic
Food production, forestry and paper, packaging, engineering, autos, chemicals, construction and building materials, diversified industrials	43	-39.031*	-2.31	-10.714		-0.79
Healthcare and pharmaceuticals	23	-10.926	-0.30	-19.589		-0.62
oil, gas, mining	10	18.212	0.65	14.775		0.53
Transport and distribution	27	5.173	0.26	40.341		2.16
Beverages, restaurant, leisure, media	33	44.940	1.59	38.219		1.25
Banks, insurance, real estate and speciality financial	34	-18.764	-1.38	-26.697		-1.69
Electronics, infotech hardware, computer software services, support services, telecommunications	40	-7.952	-0.33	-132.638*		-3.23
Retailing, stores, household goods	22	38.773	1.33	63.881		3.11
All	232	-0.982	-0.81	-13.883		-1.17

Table shows buy and hold abnormal returns (BHARs) from the sample of 232 IPOs sorted by market sector. The two columns on the right display BHARs calculated using FTSE sector indices (as opposed to the FTSE Allshare index on the left). Critical values for the skew-adjusted bootstrapped t-statistic come from the standard normal distribution, (asterisk indicates that a value weighted BHAR is significantly below zero in a one-sided hypothesis test at 5%)

The two columns on the right of Table 3.7 show that controlling for industry sector greatly changes the results on abnormal performance. Under this model the sample of IPOs delivered negative abnormal performance of 13.88%, which is greater in absolute terms than the (positive) initial returns but not statistically different from zero, and still substantially lower in absolute terms than the underperformance identified in earlier research using earlier datasets. The remarkable feature of the results is that IPOs issued in the electronics, information technology hardware,... etc sector are now seen to be dire performers. The average abnormal performance of -132.6% over 36 months is highly statistically significant and is particularly intriguing in view of the results in Table 3.4 above on initial returns, which showed that this sector experienced by far the highest initial returns. As noted above, technology related stocks performed exceptionally well

during the later part of the event period but the results in Table 3.7 show that established firms delivered far better returns than IPOs. Those who believe in investor “fads” may be able to argue that since in the time period from which the present sample arises IPOs for firms in information technology businesses were the flavour of the month, IPOs from this sector could be offered and successfully sold at unrealistically high prices. Even then, irrational investors chased the shares in early open market trading and their initial returns were consequently high. In the long-run however, initial optimism proved unjustified and the shares lost value compared to older more established information technology businesses.

3.5 Conclusion

The abnormal returns from a new sample of UK IPOs have been analysed. The choice of dataset was guided by a desire to use as recent data as possible, and to use data from the UK. These constraints led to a smaller sample than has generally been investigated in existing research. Although the initial abnormal returns reported conform closely with existing research, the long-term abnormal returns do not. Results on long-run performance are model dependent and also depend on whether equal-weighted or value-weighted *BHARs* are presented, but the benchmark calculations yield an equal weighted *BHAR* of almost exactly zero. Value-weighted *BHARs* and *BHARs* which control for industry sector are negative but not significantly different from zero. Ibbotson and Ritter (1997,

page 1007) suggest that since the long-run holding periods which are investigated must overlap, and since the number of independent observations is therefore limited, the evidence on long run returns must be considered tentative and must be treated with caution. The results reported here, and the econometric problems discussed in the previous chapter, would appear to corroborate this conclusion. Recognizing the drawback of low power emphasized by Loughran and Ritter (2000), we have not implemented any of the calendar time techniques discussed in Chapter 2. However, this would clearly have been an important step had our results indicated consistent underperformance for IPOs.

Some support for theories of irrational behaviour in stock markets may perhaps be taken from the results of dividing the sample by industry sector. IPOs in the information technology related sub-sector yielded by far the highest initial return on their first day of trading, but, especially when controlling for industry sector in the measurement of abnormal return, long term returns were extremely poor. During the period spanned by the sample, (early 1990s), it is easy to imagine that this sector could have received the most speculative interest from unconventional (noise) traders. Even if it is the case that noise traders exerted an identifiable influence on observed market prices in one sector over one period, however, the results presented in this paper do not suggest that efficient markets anomalies are pervasive with respect to IPOs.

Chapter 4. Further Features of Long-run IPO Performance

4.1 Introduction

The empirical results presented in the last chapter did not corroborate findings of existing research into long run IPO performance. Little abnormal performance was identified in the new sample and under the different techniques employed. This may support the observations of Lyon Barber and Tsai that measuring abnormal return is treacherous etc and that results must be treated with great caution but before we accept this type of conclusion we look in this chapter at other routes of investigation.

Section 2 of the chapter considers the theoretical ideas which have been advanced to explain long-run IPO underperformance and focuses on a particularly interesting idea due to Miller (1977) which is based on divergence of opinion. Miller's intuitive argument has been formalized recently by Morris (1996). This idea motivates an extension of the previous chapter's empirical investigation which is presented in Section 3 using the same sample of IPOs employed above. It is suggested that some aspects of the data do not square with the Morris model. Specifically, the abnormal performance of IPOs in the sample at hand seems to be delivered by way of low dividends rather than by poor price performance. In the next chapter an alternative model is developed which accommodates the unusual dividend process of IPOs.

4.2 Existing Theoretical Research

There is no room in mainstream portfolio theory for an argument in which the market price of a stock is simply wrong, and remains wrong over a period of years. In the presence of rational agents and in the absence of liquidity constraints or short-sales

constraints, arbitrage ought to ensure that this can never occur. Theoretical explanations for the behaviour of IPOs therefore have to reject some fundamental tenets. Shiller (1984) stressed the difference between prices which are completely random and those which are “nearly random”. In view of the literature which identifies efficient markets anomalies he argues that observed market prices are nearly random. He presents an early noise trader model in which prices are determined in part by irrational traders acting on fads and fashions, noting that these fads could be driven by anything. Starting with a demand function which contains an extra variable representing the fad Shiller uses a present value approach to arrive at an expression for price which, apart from the fad variable, is instantly recognizable as the standard present value model. He points out that depending on the specification of the fad variable the model can easily fit data which is nearly but not completely random. Later, turning to the market for IPOs, Shiller (1990) argues that investment bankers managing the floatation of an IPO can sometimes *create* demand from irrational investors. This in turn allows them to sell the IPO at a higher price, from which point it will gradually underperform. This was described by Shiller as the “Impresario Hypothesis”.

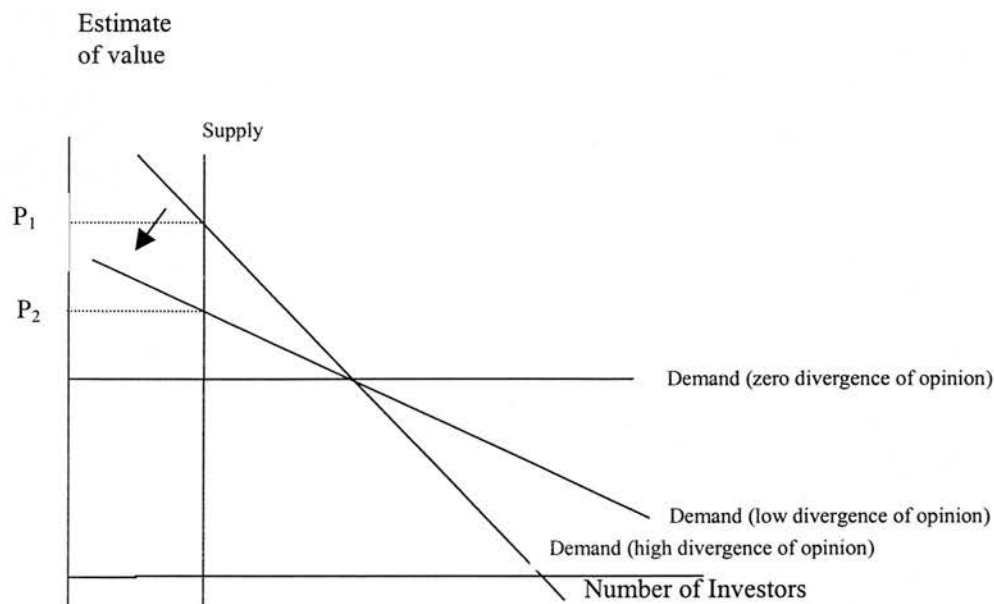
Without saying anything about why changes in investor sentiment occur, Loughran and Ritter (1995) have argued that as long as they do occur, selling equity becomes a more attractive financing option for firms at times when investor sentiment is bullish towards equities. Most IPOs arrive, therefore, at times when the market is bullish and prepared to overpay for equity, and from this point of issuance their long-run performance may be, on average, poor. Of course the performance of stocks against which the IPO is compared during the measurement of abnormal performance would also be poor so we would want to know why investors are prepared to overpay particularly for IPOs in these circumstances.

Theories which rely on the irrational behaviour of investors are anathema to many economists. The difficulty with such work lies in the representation of irrationality: a particular form of irrationality (especially a formal representation of a decision making process) can often appear *ad hoc*. In the present case where we question whether long-run IPO performance may be attributable to some aspect of irrational investor behaviour there is a further difficulty in so far as a complete theory would explain how an IPO might evolve from being an irrationally priced new issue into, after a number of years, a rationally priced ordinary stock.

Approaching the puzzle from a different viewpoint, Miller (1977) argues that if agents have *heterogeneous expectations* then the observed market price of a stock will tend to be higher than the mean valuation of (potential) investors. He argues that with heterogeneous expectations the demand curve for a risky security will be downward sloping, reflecting the willingness to pay of different investors, (see figure 1). Supply tends to be fixed. In the absence of heterogeneous expectations demand would be virtually horizontal at the (perceived) fair value of the security. As expectations become more heterogeneous, the demand curve becomes more clearly downward sloping. If supply “can be absorbed by a minority of potential purchasers (as is typically the case)”⁵⁹, the supply curve will be vertical towards the left of a supply and demand model and equilibrium price will be higher than the mean valuation of potential investors. Under these circumstances a decrease in the heterogeneity of expectations flattens the demand curve and lowers the market price (relative to its previous level).

⁵⁹ Miller, E.M., (1977), page 1153.

Figure 4.1.



Miller (1997) model of supply and demand for a security under different degrees of divergence of opinion. Diagram is drawn under the assumption that all investors have the same funds available for investment.

This mechanism will provide an instant explanation for the long-run underperformance of IPOs if only a reason can be found for why heterogeneity of expectations over the future prospects of IPOs should decrease over time. Several considerations come to mind: for example, as time goes on investors should be able to see results from new investment projects or from new areas of research, and with the higher requirements for disclosure of accounting and financial details imposed more should become known about the company after it has been listed on an official exchange.

Miller's argument has been formalized recently by Morris (1996) who models the process through which agents learn about the value of an IPO. The Miller / Morris theory of divergence of opinion has motivated some further empirical investigation of the IPO dataset at hand so we briefly summarize the Morris model here:

The Morris (1996) model of learning about the value of an IPO

There is a stock which in any particular period will pay a dividend of \$1, or will pay nothing. The dividend process is i.i.d. and the probability of a dividend in any particular period is θ . (Risk neutral) traders know that the process is i.i.d. but they do not know θ . (Their task is therefore to learn about the true value of θ), They have possibly heterogeneous prior beliefs about θ which for trader j can be represented by the density function π_j defined on the interval $[0,1]$ and they learn over time. They may buy the risky asset or a riskless asset. They may sell the riskless asset short but, importantly, they may not sell this asset short.

Valuing the asset:

If trader j 's point estimate of θ is $\hat{\theta}_j$ and the risk-free rate is r then (the risk-neutral) trader j will value the asset at $\hat{\theta}_j / r$ which is the discounted sum of expected future dividends⁶⁰. Thus, ignoring the constant, $\hat{\theta}_j$ reflects trader j 's (point) estimate of fundamental value.

Learning about θ from history

If trader j has observed a history of t periods in which s dividends have been paid then he/she has a posterior density over θ given by

$$\xi_j(\theta|s,t) = \frac{\theta^s (1-\theta)^{t-s} \pi_j(\theta)}{\int_{\zeta=0}^1 \zeta^s (1-\zeta)^{t-s} \pi_j(\zeta) d\zeta}$$

which can be conveniently understood as Bayes formula applied to a continuous variable. Trader j 's point estimate of θ (after history (s,t)) will be the mean of this posterior density:

$$\mu_j(s,t) = \int_{\theta=0}^1 \theta \xi_j(\theta|s,t) d\theta$$

After long enough history, all traders' estimates $\mu_j(s,t)$ will converge towards the true value θ .

For a specific prior density, the above expression can be solved in terms of s and t so that the convergence of $\mu_j(s,t)$ towards θ can be studied explicitly for any given history.

Market Pricing / Equilibrium

$p(s,t,r)$ is the price after a history of s dividends paid in t periods and a discount rate of r .

$\mu_*(s,t)$ is the most optimistic (highest) of all trader's valuations.

$$P(s,t,r) = \frac{1}{1+r} [\mu_*(s,t) \{1 + P(s+1,t+1,r)\} + (1 - \mu_*(s,t)) P(s,t+1,r)]$$

must be satisfied in equilibrium.

Ie, the price must be equal to the expected discounted return from holding it to the next period. If it were higher, then no-one would buy it; if it were lower, the (risk neutral)

⁶⁰ A familiar expression for the fundamental value of a stock (assuming constant expected

investor with the highest valuation would demand infinite quantities. This condition allows Morris to generate the central theorem of the paper:

Theorem

If there exists an “optimist” (a trader whose valuation is always the highest under any history of dividends), the price will always be equal to the optimist’s valuation. If there is no optimist, the price will be strictly higher than any trader’s valuation (ie a speculative premium will exist). In either case all traders’ valuations and the market price will tend to θ the true fundamental value as time goes on.

In other words when there is sufficient heterogeneity of prior expectations such that no “optimist” exists, the prevailing price will contain a speculative premium which makes it higher than any trader’s estimate of fundamental value. This occurs because the owner of the stock (the trader whose estimate of θ is the highest) calculates expected return from holding the asset one more period not just from the expected dividend but from the chance that the next period’s dividend will cause another trader’s valuation (based on history but also on prior beliefs over the distribution of θ) to become higher than his own thereby allowing the asset to be sold for a profit. As time passes and a longer history of dividends is revealed, the significance of prior expectations is diminished and the speculative premium gets smaller. This factor, and the longer history of dividends, allows the price to converge towards fundamental value.

Price is not necessarily higher than fundamental value but a speculative premium can exist and where it exists it always diminishes over time. Compared to an older asset

$$\text{return) } FV_t = \sum_{k=1}^{\infty} \frac{E_t D_{t+k}}{(1+r)^k} \text{ implies in this case } FV_t = \sum_{k=1}^{\infty} \frac{\hat{\theta}_t}{(1+r)^k} = \frac{\hat{\theta}}{r}.$$

with a long dividend history (and no significant speculative premium), the erosion of the speculative premium must generate, on average, underperformance for the newer asset during the learning process.

The learning of agents is fully rational: the only thing which diverges from standard asset pricing theory is a lack of information (and hence a heterogeneity of expectations) about the true value of an asset at the outset. This is the most complete theoretical explanation for long-run underperformance of IPOs and it is this theory which motivates the empirical investigation in Section 4.3.

4.3 Decomposition of Returns

The asset analysed in the Morris model is a stylized version of a common stock. The probability of a dividend payment in any particular period is fixed by assumption, and underperformance emerges through the erosion of a speculative premium component of the price. A priori, this pattern of price behaviour may be quite unlikely: it may be interesting to find out whether IPOs do in fact underperform due to poor price performance rather than, as may seem more likely, because they deliver lower dividends. After all, one of the main motivations for a stock market flotation is to raise money for investment, and firms with good investment opportunities are presumably more likely to retain earnings rather than pay them as dividends. In this section we attempt to decompose the returns of real IPOs into a price effect and a dividend effect: if the underperformance of such firms is driven by poor dividend payments rather than weak prices, this may be construed as evidence which does not support the theoretical argument above. Or, at least, it may prompt us to look again at the Morris model to see whether it can accommodate such an effect.

4.3.1 Data

The dataset employed in this section is an updated and expanded version of that which was collected for the analysis of the previous chapter. It covers IPOs issued on UK Stock markets between 1990Q2-1995Q1. For the present study the dataset has been extended to include price and dividend information and updated to include data up to and including November 1999.

As in the previous chapter, rights issues, seasoned equity issues, investment trust issues and government privatisations have been excluded from the sample. For all other issues, an attempt was made to collect returns information using Datastream. Datastream was also the source of all FTSE Allshare index data employed in the analysis. All of these indices are available at daily frequency. They incorporate all dividends on the appropriate dates and are adjusted for scrip issues, rights issues and other recapitalisations. Datastream also offers total returns indices derived from the corresponding FTSE Actuaries indices. These are compiled under the assumption that all dividends are reinvested in the underlying stock on the dividend payment day and are very helpful in the construction of the performance measures set out below.

Ideally, all other IPOs would be included, (including those on what was then called the Unlisted Securities Market), but Datastream does not carry researchable historic price information for all listed firms. Initially, data was collected successfully for 232 out of 288 IPOs fitting the criteria set out above; the equity raised in the 232 included offerings was £10.137bn which is 91.36% of the total amount raised in the 288 offerings. A further 14 observations have been lost in the present study (which requires extra data), because data on firms which have merged or been taken over is erased from

Datastream after a certain length of time. Thus all the calculations in the present study are based on a sample of 214 observations. The equity raised in these 214 offerings was £9.699bn.

As mentioned above, the root of the missing data problem is that many IPOs are extremely small. Frequently floated by way of placement to a small number of specialist institutional investors, such issues can be highly illiquid after flotation and, in fact, very rarely traded. When such a stock doesn't trade for days on end, a daily closing price series, even if available, may not be particularly helpful.

Table 4.1 presents summary statistics on the dataset. IPO issuance clearly varies with the economic cycle: issuance is low in the early years of the sample, a period in which the UK economy was close to, or in, recession. Stock market valuations at that time were relatively depressed. Table 2 contains the same data, sorted by market sector rather than by year of flotation. Compared with earlier datasets used in existing research, Table 2 illustrates an increase in information technology related issues and a decline in the importance of extractive industries.

Table 4.1: Summary Statistics

Year	Num ber of IPOs	total funds raised (£m)	average funds raised per IPO (£m)	total market value on flotation (£m)
1990-91	8	23.922	2.99	1728.158
1991-92	11	645.609	58.692	1587.053
1992-93	27	1597.208	59.156	4168.462
1993-94	93	3733.18	40.142	7623.93
1994-95	79	3699.413	46.828	10645.08
1990-95	218	9699.332	44.492	25752.683

Table 4.2: Summary Statistics by Industry

Market Sector	number of IPOs	total funds raised (£m)	Average funds raised per IPO (£m)	total market value on flotation (£m)
healthcare and pharmaceuticals	22	481.462	21.885	1452.912
food production, forestry and paper, packaging, engineering, autos, chemicals, construction and building materials, diversified industrials	40	2294.341	57.359	4992.478
oil, gas, mining	9	151.487	16.832	371.445
transport and distribution	26	720.654	27.717	1441.350
beverages, restaurant, leisure, media	31	1715.235	55.330	6124.251
banks, insurance, real estate and speciality financial	30	1237.625	41.254	3489.422
electronics, infotech hardware, computer software services, support services, telecommunications	38	1501.158	39.504	5048.345
retailing, stores, household goods	22	1597.370	72.608	2832.480
All	218	9699.332	44.492	25752.683

4.3.2 Methodology

As emphasized frequently above, the analysis of abnormal returns is a difficult business: in order to make statements about abnormal returns, it is first necessary to have some notion of normal returns (ie expected returns), and a model is required. The decomposition of abnormal return into two components, an abnormal price performance and an abnormal dividend performance, is a relatively straightforward extension of this difficult activity. We proceed firstly by calculating abnormal returns with reference to the market adjusted returns model as in Chapter 3 above. Then price data (exclusive of dividends) is analysed under an analogous procedure. This second stage leads to an estimate of abnormal price performance across the sample. Finally, abnormal dividend

performance is identified as the difference between abnormal returns performance and abnormal price performance.

The market adjusted returns model measures abnormal return in a particularly straightforward way:

$$(1) ar_{it} = r_{it} - r_{mt}$$

Abnormal return is the raw return on the IPO minus the return on the market for any particular holding period. Usually, the return on the market, $r_{m,t}$, will be measured by a broad share price index such as the FTSE Allshare index.

Abnormal price performance is calculated analogously:

$$(1a) ap_{it} = p_{it} - p_{mt}$$

As stated above this approach does not reflect conventional portfolio theory as no cross-sectional variation in expected return is permitted. We justify the procedure in the same way as before: all models are wrong and there is evidence (surveyed in Chapter 2 above) which suggests that the market adjusted returns model can lead to test statistics which are least badly affected by the so-called “bad model problem”

We continue with the \overline{BHAR} (buy and hold abnormal return) as our measure of long-run abnormal performance. Unlike the most popular alternative measure, the cumulative average abnormal return (CAAR), the \overline{BHAR} does not imply monthly portfolio rebalancing and conforms more closely to an ordinary understanding of a *long-run* return. In addition, although sources of test statistic misspecification exist among all

measures of long-run return, the skew-adjusted and bootstrapped statistics applied to \overline{BHAR} s below are among the most robust.

The \overline{BHAR} calculation is discussed in Chapter 3 above but the formulae are restated here for convenience:

$$BHAR_{iT} = \prod_{t=1}^T (1 + r_{it}) - \prod_{t=1}^T (1 + r_{mt})$$

$$\overline{BHAR}_T = \frac{1}{n} \sum_{i=1}^n BHAR_{iT}$$

Abnormal price performance is calculated analogously:

$$APP_{Ti} = \frac{P_{Ti}}{P_{0i}} - \frac{P_{Tm}}{P_{0m}}$$

$$\overline{APP}_T = \frac{1}{n} \sum_{i=1}^n APP_i$$

where APP_{Ti} and \overline{APP}_T are, respectively, the T period abnormal price performance of an IPO firm i and the average T period abnormal price performance of a sample of n IPO firms.

When testing for long-run abnormal returns there are potentially several sources of test statistic misspecification. One such problem may be the “bad model problem” discussed above. In tests which focus on the \overline{BHAR} another important problem will be skewness. In view of these and other issues we continue to present the skew-adjusted and bootstrapped version of the traditional t statistic as recommended in Chapter 2 and employed in Chapter 3. The skew-adjusted t statistic developed by Johnson (1978) can be written as

$$(8) \quad t_{sa} = t + \hat{\gamma} \left(\frac{1}{6n} + \frac{\overline{BHAR}_T^2}{3\hat{\sigma}^2} \right) \sqrt{n}$$

where t and $\hat{\sigma}$ are the traditional t statistic and sample standard deviation of the $BHARs$, and $\hat{\gamma}$ is the sample coefficient of skewness given by

$$\hat{\gamma} = \frac{1}{\hat{\sigma}^3} \cdot \frac{1}{n} \sum_i^n (BHAR_{iT} - \overline{BHAR_T})^3.$$

Sutton (1993) showed that bootstrapping techniques can improve the performance of t_{sa} yet further⁶¹. The t-statistics presented below have been obtained using the normal approximation method of bootstrap resampling: the details of the procedure are set out in Section 3.3 of Chapter 3 above.

4.3.3 Results

\overline{BHAR} s in Table 4.3 below are measured from the closing price on the first day of trading until the closing price on their 1st, 2nd, ..., 5th anniversaries. The starting point for long-run returns measurement is the closing price on the first day of trading rather than the issue price because there is undisputed empirical evidence which indicates that the 1 day return on IPOs is abnormally *positive*⁶². There are many theoretical arguments, mostly based on asymmetries of information, for why this initial return may be justified⁶³.

Not all of the 219 sample firms survived their first 5 years of trading, and indeed the 23 most recent issues in the sample were still less than 5 yrs old at the time when the

⁶¹ Essentially, bootstrap resampling involves treating the original sample as a population, and resampling from it with replacement.

⁶² Ibbotson and Ritter (1995) present a comprehensive survey of this subject in which they state that “The underpricing phenomenon exists in every nation with a stock market...” and their Table 1 shows the average 1 day return in empirical studies from each of 27 countries ranging from 4.2% in France to 80.3% in Malaysia

⁶³ Ibbotson and Ritter (1995) list 11 existing hypotheses which may explain the underpricing of IPOs

following calculations were completed. For those firms which have not survived, the buy and hold abnormal return from first day's closing until delisting day's closing has been measured and this return is used with an equal weight in the the \overline{BHAR} calculations⁶⁴. In the 5yr calculations this same treatment was applied to surviving firms which simply hadn't existed on the stock market for the full 5 years. The number of firms still trading in their original form on each anniversary is reported in Table 4.3.

Table 4.3 Decomposition of Abnormal Returns on IPOS 1,2,3,4 and 5 Years After Issuance

	Number of firms still trading in original form	Buy and Hold Abnormal Return (\overline{BHAR}_T)	Abnormal performance price (\overline{APP}_T)	Abnormal dividend performance	Skew (of the bhar)
1yr	219	0.019 (0.64)	0.042 (1.33)	-0.023 (-3.83)*	2.37
2yr	212	0.078 (1.42)	0.120 (2.29)*	-0.043 (-10.62)*	1.86
3yr	203	0.004 (0.08)	0.075 (0.90)	-0.070 (-6.31)*	3.00
4yr	183	-0.123 (-1.00)	-0.020 (-0.16)	-0.103 (-2.66)*	2.28
5yr	142	-0.170 (-1.07)	-0.038 (-0.245)	-0.132 (-3.64)*	2.61

Figures in parentheses are bootstrapped skew-adjusted t statistics as described in Section 3.2 above. As such critical values come from the standard normal density. Asterisk indicates that the variable is significantly different from zero at the 5% level.

Table 4.3 suggests strongly that the negative abnormal performance which exists for the IPOs in this sample is delivered more by way of a dividend shortfall than by a poor price performance. Over each of the first three years abnormal price performance is actually positive, although only over the two year period is the abnormal price performance statistically significant.

Abnormal dividend performance is, on the other hand, significantly negative over all time periods. There is strong evidence on the basis of this data that IPOs deliver lower than normal dividends.

⁶⁴ This seems perfectly reasonable in the cases where the delisting is due to takeover or merger etc, but unrealistic when the delisting is due to suspension pending bankruptcy. I have been unable to ascertain exactly the circumstances of all delistings in the sample. When I recalculate results assigning -100% raw return to the 5 firms I believe delisted pending bankruptcy, the \overline{BHAR} s are very little changed because these firms already display low raw returns)

This finding is not surprising. As stated above, firms come to the stock market as IPOs because they want money, perhaps because they have big investment ideas. If the investment opportunities which a firm faces are good and continue to be good, then we might expect the firm to retain earnings in early years and to continue to invest. In this case we should also expect positive abnormal price performance, and there may be some evidence of such a pattern in the data at hand.

We investigate the results of this experiment more closely in Tables 4.4 and 4.5. Table 4.4 breaks the data down by calendar year showing average abnormal dividend performance after 3yrs and after 5yrs. Inevitably the sub-sample sizes become rather small but the pattern of negative abnormal dividends appears quite consistent across the years. In Table 4.5 we take the same data and sort by market sector. A certain amount of cross-sector variation in average abnormal dividends is revealed, although the same caveat about the size of the sub-sample applies. Negative abnormal dividends appear to be most prevalent in 3 industry sectors: healthcare and pharmaceuticals, oil gas and mining, and electronics and IT. This finding is certainly very plausible: firms involved in pharmaceutical research, oil exploration, and certain computer software related businesses would be among those most likely to be able to float on the stock market without strong expectation of immediate profits with which to pay out dividends. In Chapter 3 we found that the healthcare and pharmaceuticals and the electronics and IT sectors delivered negative long-run abnormal returns. Although these results were not statistically significant Table 4.5 suggests that any abnormal performance appears to be delivered though an abnormally low dividend. The oil, gas and mining sector in which we find positive (but not statistically significant) abnormal performance coupled with

negative abnormal dividends appears odd but this result is based on a particularly small sub-sample.

Table 4.4 3 Year and 5 Year Abnormal dividend performance by Financial Year of Issuance

Year	N	Average 3yr Abnormal Dividend Performance	N (3yr)	Average 5 yr Abnormal Dividend Performance
1990-91	7	-0.114 (-3.20)*	7	-0.226 (-2.51)*
1991-92	11	-0.029 (-0.92)	11	0.045 (0.17)
1992-93	25	-0.098 (-7.73)*	19	-0.229 (-3.50)*
1993-94	85	-0.072 (-5.72)*	65	-0.187 (-3.69)*
1994-95	75	-0.064 (-3.87)*	40	-0.110 (-1.78)*
Entire Sample	203	-0.070 (-6.35)*	142	-0.132 (-3.80)*

Table shows abnormal dividend performance after 3yr's and after 5yr's trading, as measured from the closing price on the first day's trading. Figures in parentheses are bootstrapped skew-adjusted t statistics as described in Section 3.2 above. As such critical values come from the standard normal density. Asterisk indicates that the variable is significantly different from zero in a one-tailed test at the 5% level of significance.

Table 4.5 3 Year and 5 Year Abnormal Dividend Performance by Market Sector

Market Sector	N (3yr)	Average Abnormal Dividend Performance 3yr	N (5yr)	Average Abnormal Dividend Performance 5yr
Healthcare and pharmaceuticals	7	-0.151 (-17.88)*	7	-0.360 (-22.88)*
Food production, forestry and paper, packaging, engineering, chemicals, construction and building materials, diversified industrials	41	-0.066 (-5.43)*	32	-0.140 (-4.69)
Oil, gas, mining	8	-0.077 (-1.64)	6	-0.327 (-8.72)*
Transport and distribution	22	-0.046 (-2.62)*	14	0.014 (0.14)
Beverages, restaurant, leisure, media	25	-0.052 (-1.84)*	12	0.016 (0.10)
Banks, insurance, real estate, and speciality finance	28	-0.036 (-1.96)*	21	-0.027 (-0.45)
Electronics, IT hardware, software, telecoms	40	-0.123 (-10.83)*	26	-0.294 (-10.53)*
Retailing, stores, household goods	20	-0.005 (-0.18)	17	-0.154 (-4.65)*
Entire Sample	203	-0.070 (-6.35)*	142	-0.132 (-3.80)*

Figures in parentheses are bootstrapped skew-adjusted t statistics. Asterisk indicates that the variable is significantly different from zero in a one-tailed test at the 5% level of significance.

Overall, the results do suggest that negative abnormal dividend performance is a characteristic of IPO returns behaviour. In the next section a theoretical model is developed which allows for heterogeneous expectations over expected IPO *returns* rather than over dividends as in Morris (1996). We will be interested to see whether the model displays any of the features of IPO behaviour identified empirically.

Chapter 5. A Learning-Based Model with Unknown Expected Return

5.1 Introduction

In the last chapter we discussed the Morris model of learning about the value of an IPO. The learning approach adopted by Morris seemed particularly appropriate for two reasons: firstly, it permitted a fuller explanation of why underperformance may take place (incomplete information and divergence of opinion permit overoptimistic agents to buy); secondly it explained how the stock might stop behaving as an underperforming IPO and evolve into an ordinary and rationally priced stock (the explanation is that learning becomes complete).

In this chapter we continue to focus on the concept of learning. We develop our own model in which agents must learn about the dividend process and expected return in order to value the stock properly. We note that our model, though developed separately and for a different purpose, shares several features with a model by Lewellin and Shanken (2002) from a different branch of financial economics literature. We compare the two models and show that in one narrow sense our own model represents an extension of the LS model whereby agents have an extra variable about which to learn. We use numerical simulations to see whether the patterns of IPO performance highlighted in various parts of the thesis are discernable in the model at hand.

Section 5.2 sets out the model and relates it LS (2002), Section 5.3 presents simulation results, and Section 5.4 concludes.

5.2 The Model

A new security pays a dividend in each period. This dividend has a random component but its mean is fixed¹:

$$(1) \quad d_t = \bar{d} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2)$$

The present value model of stock prices (PVM) yields a long-run equilibrium price for the security. (The PVM expresses the price of an asset as the discounted sum of

expected future cash flows, ie. (assuming constant discount rate r):

$$(2) \quad p_t = E_t \left[\sum_{i=1}^{\infty} \left(\frac{d_{t+i}}{(1+r)^i} \right) \right]$$

In the present case the expression above can be simplified greatly since expected dividends are a constant \bar{d} :

¹ Even the fixed mean dividend process is unrealistic and its consequence is that complete learning eventually takes place. if we were to allow the dividend process to change over time learning would never cease and the key assumption in relation to IPO performance is that *more* learning takes place at the outset.

$$(3) \quad p_t = \frac{\bar{d}}{1+r} \left(1 + \frac{1}{1+r} + \left(\frac{1}{1+r} \right)^2 + \dots \right) = \frac{\bar{d}}{r}$$

In other words in the long-run price is constant and the random realizations of the dividend process are all that drive realized returns. At the outset, however, the price is not likely to be fixed as the parameter values \bar{d} and r are not known. In order to price the security investors generate estimates of \bar{d} and r using the only information available which is the dividend realizations as they arrive.

Investors know that the discount rate r is determined by the CAPM relationship

$$(4) \quad r = r_f + \beta(r_m - r_f)$$

where r_f , r_m , and β are, respectively, the risk free rate of return, the expected return on the market portfolio, and the CAPM beta reflecting covariance of the asset's return with the market portfolio.

Note that the appropriate discount rate *is* determined according to the CAPM if investors act rationally to apply a mean-variance decision rule. Since all variables are normally distributed this would be the case for example if investors are expected utility maximizers, have rational expectations and the various other assumptions re the absence of transaction costs are satisfied.

If we substitute the long-run equilibrium price (from (3)) into the returns identity

$$(5) \quad r_t = \frac{p_{t+1} - p_t + d_{t+1}}{p_t}$$

we have

$$(6) \quad r_t = \frac{d_{t+1}}{\bar{d}} r$$

which shows again that in the long-run dividends are all that drive realized returns

We make assumptions about the other assets which are largely analogous to those made about the new security. Hence the market portfolio has constant expected return r_m and the market portfolio dividend process is

$$(7) \quad d_{m,t} = \bar{d}_m + \varepsilon_{m,t} \quad \varepsilon_{m,t} \sim N(0, \sigma_{\varepsilon_m}^2)$$

Unlike \bar{d} and r , however, \bar{d}_m and r_m are known (learning has already taken place). The equation for market returns analogous to equation (6) is

$$(8) \quad r_{m,t} = \frac{d_{m,t+1}}{\bar{d}_m} r_m$$

This equation applies to realized returns directly from the outset since the parameter values are known, whereas equation (6) is an equilibrium relationship which would apply after learning has taken place.

From the outset, agents price the new security by using estimates of \bar{d} and r in equation (3).

They estimate \bar{d} as follows:

$$(9) \quad \hat{\bar{d}} = \frac{1}{t} \sum_{i=0}^{t-1} d_{t-i}$$

The estimation of r is more complicated. From the definition of β together with equations (6) and (8) we have, in equilibrium, the following:

$$(10) \quad \beta \equiv \frac{\text{cov}(r_t, r_{m,t})}{\text{var}(r_{m,t})} = \frac{\text{cov}(d_t, d_{m,t}) \frac{r}{\bar{d}} \frac{r_m}{\bar{d}_m}}{\sigma_{d_m}^2 \left(\frac{r_m}{\bar{d}_m} \right)^2}$$

$$(11) \quad = \frac{\text{cov}(d_t, d_{m,t})}{\sigma_{d_m}^2} \left(\frac{r \bar{d}_m}{r_m \bar{d}} \right) = \beta^* \left(\frac{r \bar{d}_m}{r_m \bar{d}} \right) \quad \text{where } \beta^* = \frac{\text{cov}(d_t, d_{m,t})}{\sigma_{d_m}^2}$$

Substituting (11) into the CAPM relationship (4) we have for expected return:

$$(12) \quad r = r_f + \beta^* \left(\frac{r \bar{d}_m}{r_m \bar{d}} \right) (r_m - r_f)$$

$$(13) \Rightarrow r = r_f \left(1 - \beta^* \left(\frac{\bar{d}_m}{r_m \bar{d}} \right) (r_m - r_f) \right)^{-1}$$

β^* can be estimated from observed dividends data using

$$(14) \quad \hat{\beta}_t^* = \frac{\frac{1}{t-1} \sum_{i=0}^{t-1} (d_{t-i} - \hat{d}_{t-1}) (d_{m,t-i} - \bar{d}_m)}{\sigma_{d_m}^2}$$

which is an unbiased estimator. Finally (14) and (9) can be used in (12) to provide an estimate of expected return, r :

$$(15) \quad \hat{r}_t = r_f \left(1 - \hat{\beta}_t^* \left(\frac{\bar{d}_m}{r_m \hat{d}_t} \right) (r_m - r_f) \right)^{-1}$$

If investors use the estimators in (9) and (15) in the PVM we have for prices at any given time t :

$$(16) \quad p_t = \frac{\hat{d}_t}{\hat{r}_t} = \frac{\hat{d}_t}{r_f} \left(1 - \hat{\beta}_t^* \left(\frac{\bar{d}_m}{r_m \hat{d}_t} \right) (r_m - r_f) \right)$$

$$= \frac{\hat{d}_t}{r_f} - \hat{\beta}_t^* \left(\frac{\bar{d}_m}{r_f r_m} \right) (r_m - r_f)$$

Equation (16) says that the observed market price is a linear function of two unbiased estimators of the parameters \bar{d} and β^* . As such, the unconditional expectation of p_t is the true fundamental value.

It also follows that

$$(17) \quad p = \frac{\bar{d}}{r_f} - \beta^* \left(\frac{\bar{d}_m}{r_f r_m} \right) (r_m - r_f)$$

is an alternative expression for the long-run equilibrium price, in which the risk premium is made more explicit.

Lewellen and Shanken (2002) have presented a learning-based model which they use to study not IPO behaviour but what they describe as the effects of “estimation risk”. The model set out in this chapter, though derived independently and for a different purpose, is similar to the starkest version of the LS (2002) model. A key difference is that in the latter model more is known about the security (less needs to be learned). Specifically agents know the correct discount rate from the outset and need only learn about \bar{d} . Lewellen and Shanken show that when the true mean of the dividend process is not known then even when agents are completely rational prices during the learning process display many of the patterns which have been observed in real markets and which it has often been claimed contradict the efficient markets hypothesis, (eg prices are predictable using past dividends and are excessively volatile). Lewellen and Shanken’s point is that although prices can be predicted by an

econometrician working with a long data series which permits an accurate estimate of the key parameter value, they are completely unpredictable (the martingale property holds) under the subjective distribution for dividends based on experience to date. These properties are also present in the model set out above. The essence of the LS (2002) paper is that various efficient markets anomalies which have been highlighted over the years should not be thought of as such since true expected returns are not known for future periods.

In order to compare the two models, it will be useful to set out the simplest version of Lewellen and Shanken's (2000) model. It is an Infinite Horizon Overlapping Generations model in which it is assumed that aggregate investor behaviour can be represented by a single representative agent. Much of the structure of the economy is the same as in the noise trader model of De Long, Shleifer, Summers and Waldman (1990), except of course that in LS (2002) there are no noise traders. The representative agent lives for two periods, investing an endowment when young in order to consume when old. There are, in the simplest case, two assets from which to choose: a risky one and a safe one. The latter has fixed price and pays a fixed known dividend of r . The price of the risky asset is variable and it pays a dividend which has a random component. In the case described here dividends are normally distributed about a fixed mean, (exactly as in equation (1) above). The risky asset may be interpreted as the aggregate stock market. The investor divides wealth between risky and safe asset in proportions which maximize a Constant Absolute Risk Aversion (CARA) expected utility function in terms of next period wealth. With

normally distributed returns this utility function leads to the following familiar objective function for the investor's portfolio choice:

$$(LS1) \quad \max_{x_t} EU = x_t E_t^s(p_{t+1} + d_{t+1}) + (1 - x_t)p_t(1 + r_f) - \gamma x_t^2 \text{var}_t^s(p_{t+1} + d_{t+1})$$

In LS(1) above γ is the risk aversion parameter and x_t is the number of units of the risky asset demanded. The superscript attached to the expectations operator (LS1)

indicates that the expectation integrates over the investor's subjective distribution of returns. The reason for this is that although LS state that the representative investor has rational expectations, certain parameter values (in the simplest case the mean of the dividend process) are not in the investor's information set².

The solution to the investor's portfolio choice problem is:

$$(LS2) \quad x_t = \frac{E_t^s(p_{t+1} + d_{t+1}) - p_t(1 + r_f)}{2\gamma \text{var}_t^s(p_{t+1} + d_{t+1})}$$

The equilibrium condition is that supply (normalized to 1) must equal demand.

Setting (LS2) equal to 1 and solving for price gives

² This, in fact, is the crux of the argument made by LS: they claim that their main results of returns predictability, excess volatility etc, are interesting because they arise in a model of complete rationality. It is the lack of knowledge about parameter values, however, which is critical for these results: often in macroeconomics the rational expectations solution for a model

$$(LS3) \quad p_t = \frac{1}{1+r_f} \left(E_t^s(p_{t+1} + d_{t+1}) - 2\gamma \text{var}_t^s(p_{t+1} + d_{t+1}) \right)$$

The full information steady state equilibrium with constant prices ($p_t = p$) can be found by substituting recursively in (LS3) for the term in expected price. This yields

$$(LS4) \quad p = \frac{\delta}{r_f} - \frac{2\gamma\sigma_d^2}{r_f}$$

where δ is the mean of the dividend process and σ_d^2 is it's variance.

The model can be extended to describe an economy with n risky assets rather than just one. In this case the analogous expression for the long-run equilibrium price vector is

$$(LS5) \quad \mathbf{p} = \frac{\delta}{r_f} - \frac{2\gamma}{r_f} \Sigma \mathbf{i}$$

where Σ is the covariance matrix of returns. Focusing on *gross* returns

$$(R_{i,t+1} = p_{i,t+1} - p_{i,t} + d_{i,t+1}),$$

CAPM states that

$$(LS6) \quad E_t \mathbf{R}_{t+1} = r_f \mathbf{p}_t + \beta (E_t R_{m,t+1} - r_f p_{m,t})$$

is one derived under the assumption of full information (common description of rational expectations: "agents know the correct model of the economy *including parameter values*").

In the steady state equilibrium constant prices prevail. As such the following expressions hold:

$$(LS7) \quad E_t \mathbf{R}_{t+1} = \delta, \quad E_t R_{m,t+1} = \delta_m$$

$$\beta = \frac{\text{cov}(\mathbf{R}_t, R_{m,t})}{\text{var}(R_{m,t})} = \frac{\text{cov}(\mathbf{d}_t, d_{m,t})}{\text{var}(d_{m,t})} = \frac{\sum \mathbf{i}}{\mathbf{i}' \sum \mathbf{i}}$$

These three expressions, together with (LS5), can be used in (LS6) to verify that CAPM holds for the cross-section of returns in the LS model.

(LS5) implies for the price of a single asset when many assets exist that

$$(LS8) \quad p_i = \frac{\delta_i}{r_f} - \frac{2\gamma}{r_f} \sum_{j=1}^n \sigma_{i,j} = \frac{\delta_i}{r_f} - \frac{2\gamma}{r_f} \text{cov}(d_i, d_m)$$

Equation (LS8) clearly bears a resemblance to the solution for the model presented above. Indeed since the dividend processes in the two models are identical and since the CAPM restriction applies in each case to the cross section of returns both models have the same long-run solution. The utility maximizing foundations of LS (2002) are, admittedly, a difference but they are of no essential consequence. The key difference between the two models is in the learning process since in the model developed here agents must learn about expected return as well as the mean of the dividend process, whereas in LS(2002) the dividends covariance matrix is part of the information set. The

learning process is discussed further below but first we discuss one insight that arises from the comparison of the long-term solutions of the two models.

For the purposes of comparison with (LS8), equation (17) is reprinted here:

$$(17) \quad p = \frac{\bar{d}}{r_f} - \beta^* \left(\frac{\bar{d}_m}{r_f r_m} \right) (r_m - r_f)$$

$$(18) \quad = \frac{\bar{d}}{r_f} - \frac{\text{cov}(d_i, d_m)}{\text{var}(d_m)} \left(\frac{\bar{d}_m}{r_f r_m} \right) (r_m - r_f)$$

Recognizing that $\frac{\bar{d}_m}{r_m}$ in (18) is simply p_m , and furthermore that $p_m r_m = R_m$ the (expected) gross return on the market, we can write the long-run solution in a third and final way:

$$(18) \quad p = \frac{\bar{d}}{r_f} - \left(\frac{1}{r_f} \right) \frac{(R_m - p_m r_f)}{\text{var}(d_m)} \text{cov}(d, d_m)$$

The only difference between (LS8) and equation (19) is that two times the risk aversion parameter in (LS8) is replaced by the middle term in the product on the right hand side of (19). The latter is the market price of risk. It is natural that in the LS model based on an expected utility maximization problem the market price of risk is determined by the degree of risk aversion of the representative agent. This is exactly what we have demonstrated. The utility maximizing framework of LS (2002) allows the risk premium to be determined endogenously, but of course this only emerges once the modeller has specified the degree of absolute risk aversion of the representative agent. By contrast, in the starker model set out here the expected return on the market is simply

imposed, but this expected return, when compared with the risk free rate, would imply a risk premium which in turn would imply a certain degree of risk aversion in an LS (2002) framework.

5.2.2 The predictability of returns and the excess volatility of prices and returns

In this section we show that the properties to which Lewellen and Shanken (2002) draw attention in their model are also observable in the model at hand. Lewellen and Shanken's main point is that if we accept that some information is not known to investors during a learning process then even when investors act as rationally as possible under such circumstances the observed pattern of prices may reveal unexploited profit opportunities to the econometrician who investigates *ex post*. This is a very important point because much of the empirical finance carried out in the 1980s and early 1990s attempts to refute the efficiency of financial markets by using such techniques to identify such patterns.

As do LS (2002), we focus on gross returns ($R_{t+1} = p_{t+1} - p_t + d_{t+1}$) in order to avoid the problems associated with taking expectations of ratios. Using the expression for prices in equation (16) we can write for (gross) returns:

$$(20) \quad R_{t+1} = \frac{1}{r_f} \left(\hat{d}_{t+1} - \hat{d}_t \right) - \left(\hat{\beta}_{t+1}^* - \hat{\beta}_t^* \right) M + d_{t+1} \quad \text{where } M = \frac{\bar{d}_m}{r_f r_m} (r_m - r_f)$$

$$(21) \quad \approx \frac{1}{r_f(t+1)} \left(d_{t+1} - \hat{d}_t \right) - \frac{1}{t} \left(\left(d_{t+1} - \hat{d}_{t+1} \right) \left(d_{m,t+1} - \bar{d}_m \right) - \hat{\beta}_t^* \right) M + d_{t+1}$$

The approximation in (21) lies in the second term on the right hand side. The first term on the right hand side is the corresponding term in (20) rearranged but the equivalent rearrangement of the second term ignores any change in the estimate of \bar{d} between period t and period $t+1$. Put another way, new data changes the estimate of β^* but it has a second order effect of changing the estimate of \bar{d} on which the estimate of β^* is based. The conditional expectation of (21) is:

$$(22) \quad E_t R_{t+1} \approx \frac{1}{r_f(t+1)} \left(\bar{d} - \hat{d}_t \right) - \frac{1}{t} \left(E_t \left(d_{t+1} - \hat{d}_{t+1} \right) \left(d_{m,t+1} - \bar{d}_m \right) - \hat{\beta}_t^* \right) M + \bar{d}$$

It appears from equation (22) that past dividends have an effect on expected future returns but the uncertainty surrounding β^* introduces an ambiguity compared with the model of LS (2000). High dividends in the past tend to depress expected future returns via the first term on the right hand side (as in LS (2002)) but they also have a positive effect on expected future returns via the second term on the right hand side.

An econometrician may use equation (22) to show that returns in this model are predictable but Lewellen and Shanken's point is that from the point of view of the investor, they certainly are not. Investors, assuming they begin with a diffuse prior over unknown parameter values, must base their estimates of \bar{d} and β^* on the data they have observed to date, i.e. they must use $\hat{\bar{d}}$ and $\hat{\beta}^*$. Consequently the first two terms on the right hand side of equation (22) do not

exist in the investor's (subjective) expectations and expected returns are simply \hat{d} . Prices therefore follow a martingale (from the investor's point of view).

It is also clear that prices in this model are excessively volatile. In the long-run equilibrium, (after learning has been completed), price is constant due to the constant-mean dividend process. During the learning process, however, prices are given by equation (16) and clearly fluctuate from one period to the next.

Economists no longer argue that predictable expected returns constitute evidence of market inefficiency *per se*. More often the discussion on predictability identifies two categories of explanation: rational and irrational. The source of predictability in the present model is what Lewellen and Shanken have called "estimation risk" and lies somewhere in the twilight zone between categories.

5.3 Unknown Expected Return.

Our own model, though similar to a basic version of LS (2002), has been developed for a different purpose. We have been investigating the abnormal performance of IPOs which has been identified to a certain extent in earlier chapters and reported elsewhere in the empirical literature, and we have been interested in learning based explanations. The essential difference between our own model and LS (2002) is that in the former expected return is not known. It therefore represents a second variable about which investors must learn. Unknown expected return may be a feature of IPOs more than of other

securities so we would like to see if the model captures any of the empirical patterns which have been noted, namely long-run underperformance delivered through an abnormally low dividend. We now use numerical simulations to investigate the properties of the model.

We simulate the model 500 times in order to investigate the effects of unknown expected return on price behaviour³. If estimated expected return is biased in early periods then the asset will exhibit abnormal performance from the point of view of the econometrician armed, ex post, with better parameter estimates. The benchmark simulations use equation (1) as the IPO dividend process with the unknown parameters \bar{d} and σ_e^2 set at 0.03 and 0.00015625 respectively. Equation (7), the market dividend process, has corresponding (but known) parameters \bar{d}_m and $\sigma_{\varepsilon_m}^2$ set at 0.02 and 0.0001. Dividend covariance (covariance of IPO dividends with market dividends) is 0.000075, the risk free rate is 0.02 and the expected return on the market portfolio is 0.025. As such, the true CAPM beta for the IPO is 0.444 (equation (11)), and the equilibrium price is 1.35.

Figures 5.1, 5.2 and 5.3 show the evolution of price, β_{OLS} and $\hat{\beta}$ through 500 observations of a single simulation. β_{OLS} is the econometrician's regression based estimate of the CAPM beta and differs from $\hat{\beta}$ as determined by agents in the model.

³ The simulations are carried out by running computer programmes written in GAUSS. The programme which generates the results in Table 3 below is presented as Appendix 2

It may be useful at this point to state precisely the calculation procedure for each of these variables:

- $p_t = \frac{\hat{d}_t}{r_f} \left(1 - \frac{\hat{\beta}_t^* \bar{d}_m}{r_m \hat{d}_t} (r_m - r_f) \right)$ which is estimated from the data as it arrives

exactly as in equation (16)

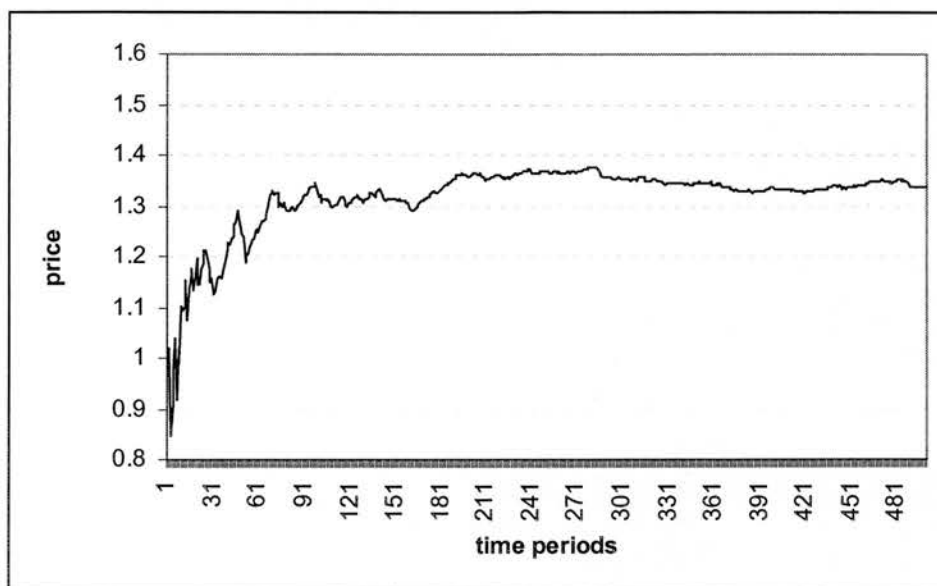
- $\beta_{OLS,t}$ is the slope coefficient from the CAPM regression of realized returns on realized market returns. This we refer to at times as the econometrician's estimate of beta.

- $\hat{\beta}_t = \hat{\beta}_t^* \left(\frac{\bar{d}_m}{r_m \hat{p}_t} \right)$ which is the estimate of the CAPM beta implied by the model (see equation 11 above).

While any such simulation is based on a unique set of data the pattern of convergence towards a constant is discernable in each of the three variables. Figures 5.1 - 5.3 are informal observations presented to give a feel for the data and while we cannot draw firm conclusions from them we get an indication from Figure 5.2 that the estimation of β_{OLS} may be problematic. In order to proceed we must analyse the behaviour of these variables on average, over repeated simulations. Table 5.1 presents the properties of these variables over 1000 simulations under the benchmark parameterisations. Each simulation

involves 490 regressions, a new regression for each period from $t=11$ onwards as one more dividend arrives⁴.

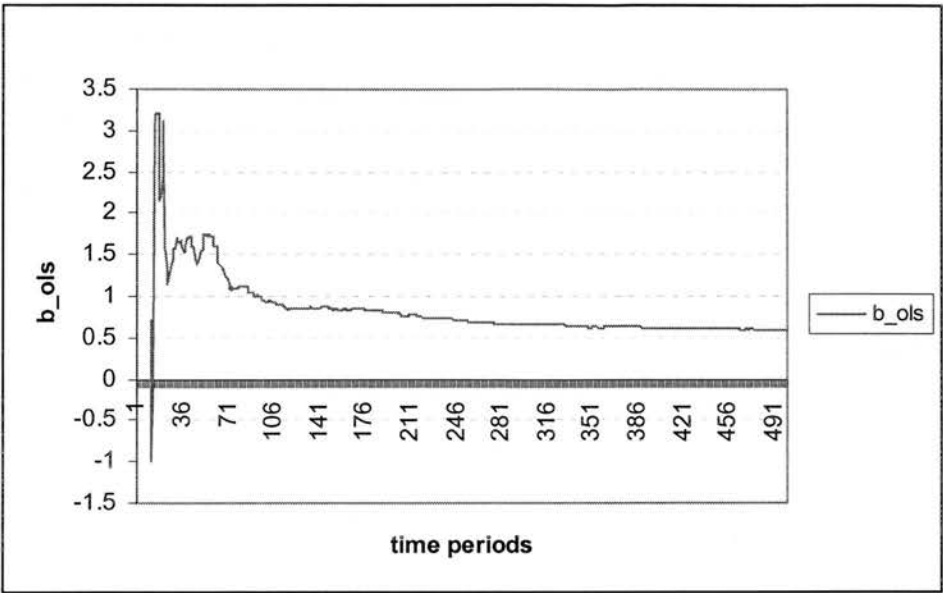
Figure 5.1: Evolution of Price during a single simulation



The graph shows the convergence of price, under the benchmark parameterisation set out above and during a single simulation, towards its equilibrium value of 1.35.

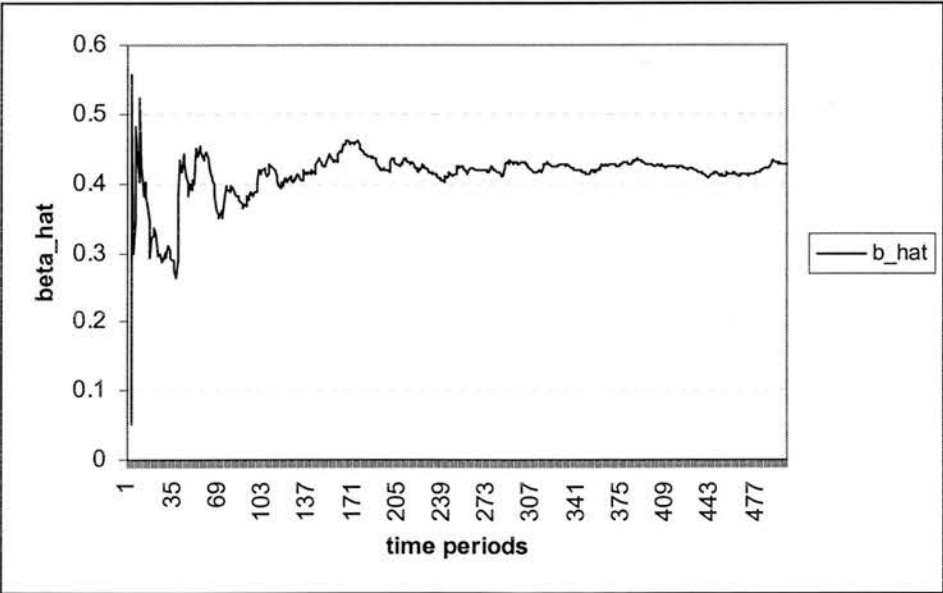
⁴ When dividends are paid semi-annually or quarterly a learning period lasting through 500 dividends is potentially dissatisfactory but we should recognize that dividend realizations are not the only factors which drive a stock price: since we only have one such factor in the model it may be reasonable to interpret the dividend more broadly as an arrival of news.

Figure 5.2: Evolution of β_{OLS} during a single simulation



The graph is based on the same data as Figure 5.1 above and shows the evolution of β_{OLS} towards its true value of 0.444. Estimation is clearly highly inaccurate during the first 100 observations and convergence is slow.

Figure 5.3: Evolution of $\hat{\beta}$ during a single simulation



The graph is based on the same data as Figures 5.1 and 5.2 above and shows the evolution of $\hat{\beta}$ towards its true value of 0.444. Estimation is not spectacular but appears far less problematic than that of β_{OLS} in Figure 5.2 above.

The results in Table 5.1 below show that the average price over the 500 simulations is close to its equilibrium level from an early stage, remarkably so in view of its variance. (An analytical expression for the variance of price has been derived in Appendix A3 and corresponds well with simulated price variance: as such, price variance in Table 5.1 is simply a function of the parameters chosen). As noted above, however, price in equation (16) is a linear function of \bar{d} , β^* and other known parameters and is therefore unbiased so our numerical results here are as expected.

On the other hand there is no reason to expect, *a priori*, any such property for β_{OLS} , the slope coefficient from ordinary least squares regression of realized returns. It appears, and this is confirmed in unreported repeat simulations using different sets of random numbers, that β_{OLS} is biased upwards compared with the corresponding true parameter value, and on average trends down towards the true parameter value over time. On the face of it this is an encouraging finding as this is the bias which is consistent with empirical findings on the long-run performance of IPOs: early estimates of expected return which are “too high” lead to a price which is “too high” and which subsequently falls as agents learn about the true parameter values. But as is evident from its variance, β_{OLS} is a very poor estimator of expected return. It is only when aggregating over 500 different data series that any relationship is discernable: at the individual level β_{OLS} is often very wild and couldn’t practically be used to estimate the CAPM Beta.

In any case, agents in the model are not interested in the CAPM Beta in its own right: they know that CAPM drives expected return and they use an estimate of dividend covariance to generate their estimate of expected return. Equation (11) relates dividend covariance to the CAPM beta and demonstrates that an estimate of the CAPM beta is implied from the estimate of dividend covariance. This implied estimate of the CAPM beta is reported in Table 5.1 as $\hat{\beta}$. A natural question to ask is whether $\hat{\beta}$ exhibits the same bias as β_{OLS} . In fact Figure 5.3 suggests and Table 5.1 confirms that under the benchmark parameters, the estimation of $\hat{\beta}$ through the structural equations of the model is clearly less problematic than the regression technique. Estimates show little sign of any bias from T=50 onwards although individual $\hat{\beta}$ estimates are still too variable for comfort and are quite inaccurate at time periods less than T=50 (this last observation is not apparent from the data presented in Table 5.1). The absence of bias in $\hat{\beta}$ ensures that agents are able, (on average), to price the stock reasonably well, a result already noted above.

Table 5.1 Benchmark Simulation Results at Various Points in Time

	T=50	T=100	T=200	T=300	T=400	T=500	T=1000
Price (sample std. dev.) (theoretical std. Dev.)	1.3477 (0.095) (0.082)	1.3527 (0.0676) (0.0580)	1.3501 (0.0467) (0.0410)	1.3512 (0.0356) (0.0334)	1.3511 (0.0309) (0.0290)	1.3521 (0.0277) (0.0259)	1.3508 (0.0189) (0.0183)
β_{OLS} (sample std. dev.)	1.3193 (0.4423)	1.0049 (0.2291)	0.7970 (0.1232)	0.7061 (0.0858)	0.6566 (0.0693)	0.6197 (0.0543)	0.5474 (0.0324)
$\hat{\beta}$ (sample std.dev.)	0.4222 (0.1347)	0.4331 (0.1010)	0.4408 (0.0715)	0.4415 (0.0573)	0.4418 (0.0491)	0.4421 (0.0449)	0.4447 (0.0309)

Results are based on 500 simulations of the model using benchmark parameter values as set out above. Snapshots of key parameters average Price, β_{OLS} and $\hat{\beta}$ over the 500 runs after 50, 100, ...1000 time periods have been taken, with variances reported in parentheses. The theoretical variance of price is derived analytically in appendix A3 and is reported for comparison. Price moves quickly towards its equilibrium value of 1.35 but $\hat{\beta}$ and especially β_{OLS} are clearly too volatile to be of use. Variance of the simulated prices is clearly of the correct order but slightly higher than that derived in the appendix: the difference is attributable to the fact that the derivation in appendix A1 assumes that \bar{d} is known rather than estimated. (reported standard deviations of price etc are standard deviations of a single price rather than the average of 500 prices)

Our main finding is that though agents in the model are able to price the security without systematic bias during the learning process, an econometrician's data-based estimate of the CAPM beta will be biased upwards. The faulty estimate would lead to an expected rate of return for the security which is "too high" and higher than the true expected rate of return. As such, a CAPM based approach to measuring abnormal performance would be expected, by incorporating this bias, to generate negative abnormal returns.

We try to learn more about the sources of the bias in β_{OLS} by varying the parameters. There are in fact only 4 free parameters in the model. They are r_f , the risk free rate, r_m , the expected return on the market portfolio, \bar{d}_m , the mean

dividend on the market portfolio and β^* , the dividend covariance parameter: changes in σ_d , the variance of IPO dividends, have no effect on equilibrium price or expected return reflecting, as they do, idiosyncratic risk; changes in σ_{d_m} , the variance of market dividends, have a direct and proportional effect on β^* and we choose to control β^* ; and \bar{d} , average IPO dividends, covaries with \bar{d}_m .

We run the same simulations as those which underlie Table 5.1 but with 81 different sets of parameter values. Table 5.2 presents the various parameter values which have been employed.

Table 5.2 Parameter Values

Parameter	Values taken		
Risk free rate (r_f)	0.	0.	0.02
Expected return on market (r_m)	0.	0.	0.03
Market dividend (\bar{d}_m)	0.	0.	0.04
Covariance parameter (β^*)	-	0.	1.75

The table presents the parameter values employed in simulations designed to identify sources of bias: 3 values for each of 4 parameters allow 81 separate simulation exercises.

These parameter values generate equilibrium prices ranging from 0.35 to 4.7, and true CAPM betas ranging from -1.14 to 0.81. As in the initial experiment described in Table 5.1, price is estimated successfully under each of the 81 parameterisations but the purpose is to look for the source of bias in β_{OLS} . To this end we regress the 81 estimation errors onto the parameter values which generated them. The regression equation is

$$(23) (\beta_{OLS} - \beta) = \alpha_0 + \alpha_1 r_f + \alpha_2 r_m + \alpha_3 \bar{d}_m + \alpha_4 \beta^* + \varepsilon$$

for which we have 81 data points each consisting of a β_{OLS} estimation error and the combination of parameter values from Table 5.2 which applied in that particular loop of the simulation programme. Equation (23) linearizes the relationship between the estimation error and the other variables in the model in an attempt to see which of the other variables exert significant influence. The regression results are presented in Table 5.3. It is clear from the results that β^* is the only important determinant of estimation error in the experiment. This suggests that β_{OLS} is more heavily upwards biased when β^* is high.

This is a significant finding in itself as β^* is obviously positively related to the true CAPM beta. It would suggest that in a study of abnormal performance in which CAPM determines expected return, stocks with high betas would be those most likely to display negative abnormal performance. For reasons discussed in Chapter 2 we have not estimated CAPM Betas in the empirical section of the thesis but it seems likely from their strong total returns performance that stocks from the telecoms and IT sectors might have the highest CAPM betas and it is interesting to note that it was these sectors which generated the most significant underperformance in the empirical exercise at the end of Chapter 3.4.2 under a completely different (and hopefully more robust) methodology.

Table 5.3. Regression of OLS Regression Error on Parameters

Variable	Coefficient	Standard error
Intercept	-0.160	0.219
Risk free rate (r_f)	-0.848	5.885
Return on Market Portfolio (r_m)	1.086	5.885
Dividend on market portfolio (\bar{d}_m)	-0.424	2.942
Covariance parameter (β^*)	0.495	0.029
Observations		81
Adjusted R Square		0.777
F		70.784

Results of estimation of Equation (23) using data from simulations. Only the covariance parameter is statistically significant

We now carry out another straightforward experiment to investigate the extent of bias in β_{OLS} . We use data from the same 81 simulations as in the experiment above and we regress β_{OLS} onto the corresponding true CAPM beta. The results, presented in Table 4, clearly confirm that β_{OLS} is heavily biased upwards.

Table 5.4 Regression of β_{OLS} onto β

Variable	Coefficient	Standard Error
Intercept	0.038	0.016
True Beta	1.402	0.032
Observations		81
Adjusted R square		0.960

The table shows results of the simple regression $\beta_{OLS} = a + b\beta_{CAPM} + e$ where the dependent variable is the estimate of the true CAPM beta which comes out of the simulation exercise above (ie each β_{OLS} observation is the average of the last estimate of β_{OLS} over 500 simulations under one of the parameter combinations available from Table 5.3 above.)

5.4 Conclusions

We have set out a model in which CAPM determines the rate of return on an asset. In the case of a new security (an IPO) investors understand the model of

the economy but must learn key parameter values, specifically expected return and average dividend, in order to price the security. We compare the model to a basic version of another model due to Lewellen and Shanken (2002), noting some similar properties and demonstrating that our model, though employed for a different purpose, can be viewed as an extension of the basic LS (2002) model.

We simulate our model using randomly generated data and find some difficulties with precision of estimates during the learning period, but we do not find systematic biases in the key parameters as estimated by agents in the model. By contrast, taking the point of view of the econometrician investigating the CAPM beta ex post on the basis of historic data, we find clear evidence of upward bias which would lead to an upwardly biased estimate of expected return. This bias is driven by β^* , the covariance parameter, and recedes slowly over time. Nevertheless, an upwardly biased estimate of expected return would generate a spurious (negative) abnormal return of the type noted in research on the long-run performance of IPOs. To reinforce the point, observed underperformance of IPOs would be entirely spurious if it were generated by the estimation problems which have come to light during the investigation of the model at hand. Rather than drawing this conclusion from a model which is still quite stark and stylised, we simply suggest that the finding supports the empirical practice of not estimating the CAPM beta and using instead some other benchmark for expected returns in the measurement of long-run abnormal security returns.

Chapter 6. Conclusion

The research presented in this thesis does not lead us to a firm conclusion on the subject of long-run IPO performance. Empirical evidence in Chapters 3 and 4 identifies a certain amount of underperformance among certain types of firm when we control for industry sector in the measurement of abnormal return, but overall the results are clearly less emphatic than those reported in earlier research. We have adopted a cautious approach from the start, analysing in detail the econometric difficulties involved in measuring and testing abnormal returns and employing as many as possible of the recommendations in our own empirical work. It is possible that the up-to-date techniques employed have served to reduce any spurious element of negative abnormal performance which may have existed in any of the earlier studies referred to in the text. What is more likely, however, is that underperformance of IPOs is simply less pronounced in the new data examined here. (As mentioned in the text many of the "Efficient Markets anomalies" have tended to become less pronounced over time). To a certain extent time will tell on this question as the subject will surely be revisited in the future using new data and yet different techniques.

The second half of the thesis has been concerned with economic theory which might explain long-run underperformance. One firm conclusion from our research is that if such systematic long-run underperformance exists then the explanation is likely to be connected with the concept of learning. It is inherently plausible that investors know less at the outset about firms which are newly listed on stock exchanges. In addition, learning-based models are unique

in being able to explain the transition of a firm from abnormally performing IPO into normally performing stock. In Chapter 5 where we develop a learning based model with unknown expected return we come to a conclusion which harks back to the econometric difficulties discussed in Chapter 2. We find that although agents within the model are able to price the new security and estimate the CAPM beta without bias, this is not how it might appear to the econometrician investigating ex post: numerical simulations reveal that the OLS estimate of the CAPM beta is biased upwards, a finding which could lead to upwardly biased estimates of expected return and, in turn, negative abnormal return during the learning period. We also find that the size of the bias varies positively with the true CAPM beta. This is particularly interesting if we believe that the hi-tech firms which displayed negative abnormal performance in the empirical study of chapter 3 are those most likely to have high CAPM betas. This is an area which would be worthy of further research.

After approaching the issue from various directions it is still not clear whether IPOs underperform in the long-run. Nevertheless, as well as contributing to the research on this specific question, it is hoped that the work in this thesis shows how the various approaches to the investigation of this or any other Efficient Markets puzzle can be put into practice.

Appendix A1. Gauss Programme for Various Skew Adjusted t-test Procedures

This programme computes the various bootstrapped skew adjusted t-statistics described in the in Chapter 2 of the text, (Section 2.3.1, pages 24-25). The Normal Approximation Bootstrap Resampling Test statistic, generated last, is the one which has been presented with empirical results throughout Chapters 3 and 4.

```
new;
t1=hsec;
output file = mean_ci.out;
output reset;

/* Parameters to be set for the bootstrap run*/;
n = 232;
n_bt = 232;
tcal = 1.98;
b = 1000;

print "n = " n;
print "b = " b;
print "tcal = " tcal;
print;
i = 1; j = 1;
prop = 0;
x_bar_bt = zeros(B,1);
sd_xb_bt = zeros(B,1);
tbb = zeros(B,1);
skewb = zeros(B,1);

load x[232,1] = bhar.asc; /*loads the 36 mnth bhar data*/;

/*Full Sample Results*/;
x_bar = meanc(x);
sd_x = stdc(x);
sd_x_bar = sd_x/sqrt(n);

/*Resampling*/;
do while i<=b;
index = ceil(rndu(n_bt,1)*n); /*set up resampling index*/;
x_b = x[index,.];
m1b = x_b-meanc(x_b);
m3b = sumc(m1b^3)/n_bt;
sd_xb = stdc(x_b);
sb = meanc(x_b)/sd_xb;
skewb = (m3b/(sd_xb^3)); /*calculating the skew statistic of a resample*/;
```

```

x_bar_b = meanc(x_b); /* calculating x-bar for a resample*/;
sd_xb_b = stdc(x_b)/sqrt(n_bt);
tb = sqrt(n_bt)*(sb+skewb*(sb^2)/3+skewb/6/n_bt);
x_bar_bt[i,1] = x_bar_b; /* setting up a vector of bootstrapped x-bars*/;
sd_xb_bt[i,1] = sd_xb_b;
tbb[i,1] = tb;
i = i+1;
endo;

/* Confidence Intervals*/;
pmlcl = x_bar - (tcal*sd_x_bar);
pmucl = x_bar + (tcal*sd_x_bar); /*parametric CI*/;
tstat = x_bar/sd_x_bar;

/*skew statistic*/;
m1 = x - meanc(x);
m3 = sumc(m1^3)/n;
skew = (m3/(sd_x^3));

/*skew adjusted t-statistic*/
s = x_bar/sd_x;
t_sa = sqrt(n)*(s+skew*(s^2)/3+skew/6/n);

sd_tbb = stdc(tbb);
z = t_sa/sd_tbb;

x_bar_bt = x_bar_bt~sd_xb_bt;
x_bar_bt = sorthc(x_bar_bt,1);
ll = round((b*0.05)/2);
ul = round((b-ll)+1);
perlcl = x_bar_bt[ll,1];
perucl = x_bar_bt[ul,1];

tbb = sorthc(tbb,1);
tbbll = round((b*0.05)/2);
tbbul = ((b-tbbll)+1);
lower_CV = tbb[tbbll,1];
upper_CV = tbb[tbbul,1];

/*printing results*/;
print "full sample x-bar value = " x_bar; format /rd 2,5;
print "full sample SD (x-bar) = " sd_x_bar; format /rd 2,5;
print "parametric t statistic = " tstat;
print;
print "bootstrap estimate of x-bar = " meanc(x_bar_bt[:,1]);
print "bootstrap estimate of SD (x-bar) = "stdc(x_bar_bt[:,1]);
print;
print "parametric CI = " pmlcl pmucl;
print "percentil CI = " perlcl perucl;
print;

```

```
print "skew adjusted t-statistic = " t_sa;
print "critical values for skew adjusted t-statistic" lower_CV upper_CV;
print;
print "Normal Approximation Bootstrap Resampling Test";
print "st dev (t_sa) = " sd_tbb;
print "z = " z;

end;
```

Appendix A2. Gauss Programme for Generating Simulation Results of Table 5.3

```
/* NAME OF PROGRAMME: table 3 */;

NEW;
output file = b_500_.asc;
output reset;

result=zeros(81,4);

/* PARAMETER SETTINGS */;
t=500;

counter=1; /*added */;
for g (1,3,1);
    for h (1,3,1);
        for l (1,3,1);
            for m (1,3,1);

                r_m=.02+g*.005; /*arranging the for loops*/;
                r_rf=0.005+h*.005;
                D_bar_m=.01+.01*l;
                cov=-1.25+m;

                sig_d_m=.0001;

                s=D_bar_m/r_m*(r_m-r_rf);

n=200;
e=10; /* no of obs excluded from regressions */
big_b=zeros(n,2);
j=1;
do while j<=n;

    /*SETTING UP EMPTY VECTORS AND INITIAL VALUES */;
    e_mkt=zeros(t,1);
    e_ipo=zeros(t,1);
    cm_earn=zeros(t,2);
    earn_bar=zeros(t,2);
    dme_m=zeros(t,1);
    dme_i=zeros(t,1);
    b_star=zeros(t,1);
    r_hat=zeros(t,1);
```

```

price=zeros(t,1);
r_real_i=zeros(t,1);
r_real_m=zeros(t,1);
b_ols=zeros(t-e-10,2); /*note smaller matrix */;
wunz=ones(t,1);
x=wunz~r_real_m-r_rf; /* the X matrix for CAPM regressions */;

e_ipo[1,1]=0.02;
e_mkt[1,1]=0.02;
cm_earn[1,]=rndn(1,2)/100+D_bar_m;
earn_bar[1,]=cm_earn[1,];
b_star[1,]=1;
r_hat[1,]=r_rf;
price[1,]=D_bar_m/r_hat[1,];
r_real_i[1,]=r_hat[1,];
r_real_m[1,]=r_hat[1,];
b_ols[1,1]=0;
b_ols[1,2]=0;
x[1,1]=1;
x[1,2]=0;

/*CALCULATING RESULTS */;
i=2;
do while i<=t;
    e_mkt[i,1]=rndn(1,1)*sqrt(sig_d_m)+D_bar_m;
    e_ipo[i,1]=.015+cov*e_mkt[i,1]+rndn(1,1)/100;
    earn=e_ipo[i,1]-e_mkt[i,1];
    cm_earn[i,]=earn+cm_earn[i-1,];
    earn_bar[i,]=cm_earn[i,]/i;
    dme_m[1:i,1]=e_mkt[1:i,1]-D_bar_m;
    dme_i[1:i,1]=e_ipo[1:i,1]-earn_bar[i,1];
    b_star[i,]=sumc(dme_i.*dme_m)/((i-1)*sig_d_m);
    r_hat[i,1]=r_rf/(1-b_star[i,1]/earn_bar[i,1]*s);
    price[i,1]=earn_bar[i,1]/r_hat[i,1];
    r_real_i[i,1]=(price[i,1]-price[i-1,1]+e_ipo[i,1])/price[i-1,1];
    r_real_m[i,1]=e_mkt[i,1]/(D_bar_m/r_m);
    x[i,2]=r_real_m[i,1]-r_rf;

    if i>e+10;

        b_int=inv(x[e+1:i,1]*x[e+1:i,1]')*x[e+1:i,1]'(r_real_i[e+1:i,1]-r_rf);

        b_ols[i-e-10,]=b_int'; /* b_ols is t-e x 2 */;
    endif;

    i=i+1;
enddo;

```



```

big_b[j,.]=b_ols[t-e-10,2]~price[t,.]; /* big_b is 200 x 2 */;

j=j+1;

endo;

result[counter,.]=meanc(big_b)'~stdc(big_b)';

counter=counter+1;

    endfor;
endfor;
endfor;
endfor;

/*RESULTS*/;

result;

hist(big_b,8);

/* GENERATING CHART OUTPUT */
x=seqa(1,1,t-e-10);
library pgraph;
_pltype={6,6};
xy(x,b_ols[.,2]);

```

Appendix A3. Derivation of variance for p_t and $\hat{\beta}_t^*$

$$\hat{\beta}_t^* = \frac{\frac{1}{t-1} \sum_{i=1}^t (d_i - \bar{d}) (d_{m,i} - \bar{d}_m)}{\sigma_{d_m}^2}$$

Firstly, looking for the variance of $(d_i - \bar{d})(d_{m,i} - \bar{d}_m)$, the product of two normal variables:

$$(1) \quad \text{var}[(d_i - \bar{d})(d_{m,i} - \bar{d}_m)] = E[(d_i - \bar{d})^2 (d_{m,i} - \bar{d}_m)^2] - \sigma_{d,d_m}^2$$

and writing $(d_i - \bar{d})$ as

$$(2) \quad (d_i - \bar{d}) = \rho \frac{\sigma_d}{\sigma_{d_m}} (d_{m,i} - \bar{d}_m) + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma_\varepsilon^2), \quad E[\varepsilon_i d_{m,i}] = 0$$

Substituting (2) into (1):

$$\begin{aligned} \text{var}[(d_i - \bar{d})(d_{m,i} - \bar{d}_m)] &= E \left[\left(\frac{\sigma_{d,d_m}^2}{\sigma_{d_m}^4} (d_{m,i} - \bar{d}_m)^2 + \varepsilon_i^2 - 2\varepsilon_i \rho \frac{\sigma_d}{\sigma_{d_m}} (d_{m,i} - \bar{d}_m) \right) (d_{m,i} - \bar{d}_m)^2 \right] - \sigma_{d,d_m}^2 \\ &= E \left[\frac{\sigma_{d,d_m}^2}{\sigma_{d_m}^4} (d_{m,i} - \bar{d}_m)^4 + \varepsilon_i^2 (d_{m,i} - \bar{d}_m)^2 \right] - \sigma_{d,d_m}^2 \\ (3) \quad &= 3\sigma_{d,d_m}^2 + \sigma_\varepsilon^2 \sigma_{d_m}^2 - \sigma_{d,d_m}^2 \end{aligned}$$

($d_{m,i}$ is normal and as such its 4th central moment is equal to $3\sigma_{d_m}^2$)

From (2) we can also say that

$$(4) \quad \text{var}(\varepsilon_i) = \sigma_d^2 + \frac{\sigma_{d,d_m}^2}{\sigma_{d_m}^2} - 2 \frac{\sigma_{d,d_m}^2}{\sigma_{d_m}^2} = \sigma_d^2 - \frac{\sigma_{d,d_m}^2}{\sigma_{d_m}^2}.$$

Substituting (4) into (3),

$$\begin{aligned} \text{var}[(d_i - \bar{d})(d_{m,i} - \bar{d}_m)] &= 2\sigma_{d,d_m}^2 + \sigma_{d_m}^2 \left(\sigma_d^2 - \frac{\sigma_{d,d_m}^2}{\sigma_{d_m}^2} \right) \\ (5) \quad &= \sigma_{d,d_m}^2 + \sigma_d^2 \sigma_{d_m}^2 \end{aligned}$$

Secondly, looking for the variance of the sum $\sum_{i=1}^t (d_i - \bar{d})(\bar{d}_{m,i} - \bar{d}_m)$:

$$(6) \quad \text{var}\left[\sum_{i=1}^n (d_i - \bar{d})(\bar{d}_{m,i} - \bar{d}_m)\right] = t(\sigma_{d,d_m}^2 + \sigma_d^2 \sigma_{d_m}^2)$$

Finally,

$$(7) \quad \text{var}(\hat{\beta}_t^*) = \frac{1}{\sigma_{d_m}^2} \frac{t}{(t-1)^2} (\sigma_{d,d_m}^2 + \sigma_d^2 \sigma_{d_m}^2)$$

For price we have

$$(8) \quad p_t = \frac{\hat{d}_t}{r_f} - \hat{\beta}_t^* \left(\frac{\bar{d}_m}{r_f r_m} \right) (r_m - r_f)$$

which is equation (16) in the text. Denoting $\left(\frac{\bar{d}_m}{r_f r_m} \right) (r_m - r_f)$, a known constant, as k , the variance of p_t is

$$(9) \quad \text{var}(p_t) = \frac{\sigma_d^2}{r_f^2 t} + k^2 \left(\frac{1}{\sigma_{d_m}^2} \frac{t}{(t-1)^2} (\sigma_{d,d_m}^2 + \sigma_d^2 \sigma_{d_m}^2) \right).$$

Equations (7) and (9) have proved useful for checking that the simulations in section X are running properly¹.

¹ between 5% and 15% too high, probably because \bar{d} is not known in simulations

References

- Barber, B.M., and Lyon, J.D. (1997): "Detecting Long-Run Abnormal Stock Performance: The Empirical Power and Specification of Test Statistics", *Journal of Financial Economics*, vol. 43, pp. 341-372
- Black, F., (1972): "Capital Market Equilibrium with Restricted Borrowing", *Journal of Business*, vol. 45, pp.444-454
- Blume, M.E., and Stambaugh, R.F., (1983): "Biases in Computed Returns: An Application to the Size Effect", *Journal of Financial Economics*, vol. 12, pp.387-404
- Brav, A., (2000) "Inference in Long-Horizon Event Studies: A Bayesian Approach with Application to Initial Public Offerings", *Journal of Finance* vol. 55, no.5.
- Brav, A., Geczy, C. and Gompers, P., (2000), "Is the Abnormal Return Following Equity Issuances Anomalous?", *Journal of Financial Economics*, vol. 56,
- Brav, A., and Gompers, P.A., (1997): "Myth or Reality? The Long-Run Underperformance of Initial Public Offerings: Evidence from Venture and Non-Venture Capital-Backed Companies", *Journal of Finance*, vol. 52, no. 5
- Brown, E., (1999): "Long-Run Performance Analysis of a New Sample of UK IPOs", Department of Economics discussion paper 99-10, University of Edinburgh
- Brown, S.J., and Warner, J.B. (1980): "Measuring Security Price Performance", *Journal of Financial Economics*, vol. 8, pp. 205-258
- Brown, S.J., and Warner, J.B. (1985): "Using Daily Stock Returns: The Case of Event Studies", *Journal of Financial Economics*, vol. 14, pp. 3-31
- De Long, J.B., Shleifer, A., Summers, L.H., and Waldmann, R.J. (1990): "Noise Trader Risk in Financial Markets", *Journal of Political Economy*, vol. 98, no. 4, pp. 703-738

Espenlaub, S., Gregory, A., and Tonks, I. (1998): "Testing the Robustness of Long Term Underperformance of UK Initial Public Offerings", LSE Financial Markets Group discussion paper 285.

Fama, E.F. (1991): "Efficient Capital Markets: II", *Journal of Finance*, vol. 46, no. 5, pp. 1575-1617

Fama, E.F. (1998): "Market efficiency, long-term returns, and behavioural finance", *Journal of Financial Economics*, vol. 49, no. 3, pp. 283-306

Fama, E., and French, K., (1992), "The Cross-Section of Expected Stock Returns", *Journal of Finance*, vol. 47, pp. 427-465.

Fama, E., and French, K., (1993), "Common risk Factors in the Returns on Stocks and Bonds", *Journal of Financial Economics*, vol. 25, pp. 23-49.

Ibbotson, R.G., (1975): "Price Performance of Common Stock New Issues", *Journal of Financial Economics*, vol. 2, pp. 235-272

Ibbotson, R.G., and Jaffe, J.F., (1975): "Hot Issues' Markets", *Journal of Finance*, vol. 30, pp.1027-1042

Ibbotson, R.G., and Ritter, J.R. (1997): "Initial Public Offerings" in Jarrow, R.A., Maksimovic, V., and Ziemba, W.T. (ed.), *Handbooks in Operations Research and Management Science*, Elsevier Science, pp. 993-1017

Johnson, N.J., (1978): "Modified *t* Tests and Confidence Intervals for Asymmetric Populations", *Journal of the American Statistical Association*, vol. 73, pp. 536-544.

Kothari, S.P., and Warner, J.B. (1997): "Measuring Long Horizon Security Returns", *Journal of Financial Economics*, vol. 43, pp. 301-339

Lee, C.M.C., Shleifer, A., and Thaler, R.H. (1991): "Investor Sentiment and the Closed-End Fund Puzzle", *Journal of Finance*, vol. 46, no. 1, pp. 75-109

Lyon, J.D., and Barber, B.M., and Tsai, C., (1998) "Improved Methods for Tests of

Long-Run Abnormal Stock Returns", *Journal of Finance*, vol. ??, no. 5??, pp. ???

Keim, D., (1989): "Trading Patterns, Bid-Ask Spreads, and Estimating Security Returns: The Case of Common Stocks at Calendar Turning Points," *Journal of Financial Economics*, vol. 25, pp. 75-97.

Kothari, S.P., and Warner, J.B. (1997): "Measuring Long Horizon Security Returns", *Journal of Financial Economics*, vol. 43, pp. 301-33

Lee, P.J., Taylor, S.L., and Walter, T.S. (1996): "Australian IPO pricing in the short and long run", *Journal of Banking and Finance*, vol. 20, no. 7, pp. 1189-1210

Levis, M. (1993): "The Long-Run Performance of Initial Public Offerings", *Financial Management*, vol. 22, pp. 26-41.

Levis, M. (1995): "Seasoned Equity Offerings and the Short and Long-run Performance of IPOs in the UK", *European Financial Management*, vol. 1, no. 2, pp. 125-146

Lewellen, J., and Shanken, J., (2002), "Learning, Asset Price Tests, and Market Efficiency", *Journal of Finance*, (forthcoming)

Loughran, T., and Ritter, J.R. (1995): "The New Issues Puzzle", *Journal of Finance*, vol. 50, pp. 23-51

Loughran, T., and Ritter, J.R. (2000): "Uniformly Least Powerful Tests of Market Efficiency", *Journal of Financial Economics*, vol. 55, pp. 361-389.

Lyon, J.D., Barber, B.M., and Tsai, C., (1998): "Improved Methods for tests of Long-Run Abnormal Stock Returns", *Journal of Finance*, vol. 54, no. 1, pp. 165-203.

Mehra, R., and Prescott, E. (1985): "The Equity Premium: A Puzzle", *Journal of Monetary Economics*, vol. 15, pp. 145-62

Miller, E.M. (1977): "Risk, Uncertainty and Divergence of Opinion", *Journal of Finance*, vol. 32, no. 4, pp. 1151-1168

Mitchell, M., and Stafford, E., (2000) "Managerial Decisions and Long-Term Stock Performance." *Journal of Business*, vol. 73, no. 3

Morris, S. (1996): "Speculative Investor Behaviour and Learning", *Quarterly Journal of Economics*, vol. 111, no. 4, pp. 1111-1133

Ritter, J.R. (1991): "The Long-Run Performance of Initial Public Offerings", *Journal of Finance*, vol. 46, pp. 3-27

Samuelson, P., (1965) : "Proof that Properly Anticipated Prices Fluctuate Randomly", *Industrial Management Review*, vol.6, pp. 41-49

Sharpe, W., (1964) : "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk", *Journal of Finance*, vol.19, pp. 425-442

Shiller, R.J. (1984): "Stock Prices and Social Dynamics", *Brookings Papers on Economic Activity*, vol. 2, pp. 457-498

Shiller, R.J., (1990): "Speculative Prices and Popular Models", *Journal of Economic Perspectives*, vol. 4, pp. 55-65

Shleifer, A., and Vishny, R.W. (1997): "The Limits of Arbitrage", *Journal of Finance*, vol. 52, no.1, pp. 35-55

Sutton, C.D., (1993): "Computer-Intensive Methods for Tests About the Mean of an Asymmetrical Distribution", *Journal of the American Statistical Association*, vol. 88, no.423, pp. 802-810